VOLUME X SEPTEMBER, 1919 NUMBER 7

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

A GROUP SCALE OF INTELLIGENCE FOR USE IN THE FIRST THREE GRADES

Studies from the Psychological Laboratory of Indiana University

LUELLA WINIFRED PRESSEY[1]

1. SPECIAL PROBLEMS INVOLVED IN THE DEVELOPMENT OF GROUP TESTS FOR YOUNG CHILDREN

The study described in the present paper began as an effort to devise a brief group scale of general intelligence for use (1) as a means by which children might be tested upon entering school and (2) as an aid in grouping these children into grades or sections, according to ability. It seemed to the writer that data of unusual interest might be gathered if such measurements of the entering classes – – the "pupil material" with which the school is to deal — might be obtained. Comparison of different schools as to the quality of their "pupil material" should be particularly interesting. However, such a scale might prove of greater value as a means of measuring the accuracy with which, in the first grades, children are selected according to ability for "fast" and "slow" sections, or are placed at entrance in their proper grade. In both these ways, then, such a scale should prove of unique value.

However, little work has thus far been done in the development of group tests, of either achievement or endowment, for use with such young children. Very early in the investigation, it appeared that work with first grade children involved certain special problems of test organization and technique, in the study of which the writer became quite as much interested as in the more specific effort to

NOTE 1—This paper is one of a series of studies being made by the Department of Psychology, Indiana University, in the field of mental tests. A later paper will deal with a scale complimentary to the examination described in the following paper, and intended for use in grades 4-8 inclusive. This second "Cross-Out" scale has already received brief notice in the Notes of this Journal.

develop a particular groups of tests. The paper should, then, be considered quite as much a study in the general method and technique necessary for work with young children, as a presentation of a particular scale particular set of methods.

As the writer analyzed the situation, the difficulties involved in the formulation of group tests for use with first grade children, were five in number. None of these difficulties appear in such prominence, if at all, in the making of tests for older children. In the first place, there were two special problems of material. (1) Older children can read and write; but no reading or writing, even of letters or numbers, can be involved in tests for young children. (2) Older children can work for some time without marked fatigue; young children cannot. It is very important that the examination be short and that the clerical labor of taking the examination be cut down to a minimum. Added to these two difficulties are three others having to do with the presentation of the tests. (3) Older children can understand and follow uninteresting and formal directions, if only they are clear. This is not true of little children. All kindergarten and first grade work is based on the principle of "spontaneous interest and participation" on the part of the children. This must be constantly kept in mind when giving tests to them. They must take an active part in the directions. In fact, their part must be the major one, the examiner merely guiding them through the examination. (4) Older children do not need a carefully integrated scale; they can easily shift from one problem to another, from one method of indicating score to another, or from one form of directions to another. Small children simply will not shift readily from one type of problem or reaction to another; there is a loss of interest and a period of general confusion after each test, before they are successfully launched on the next – – if, in fact, they ever do get started. Only by the most compact organization can the childrens' interest and attention be sustained from one test to the next. (5) Finally, older children will pay attention and do reliable work even if they are not particularly interested in the examination; habits of work and school discipline will keep them at their task. Primary children refuse point blank to work unless they are interested. There is here, in fact, a fundamental problem of motivation, which the writer considers of the greatest importance. Primary children are not accustomed to rigid discipline. They have not yet learned to work and simply cannot be

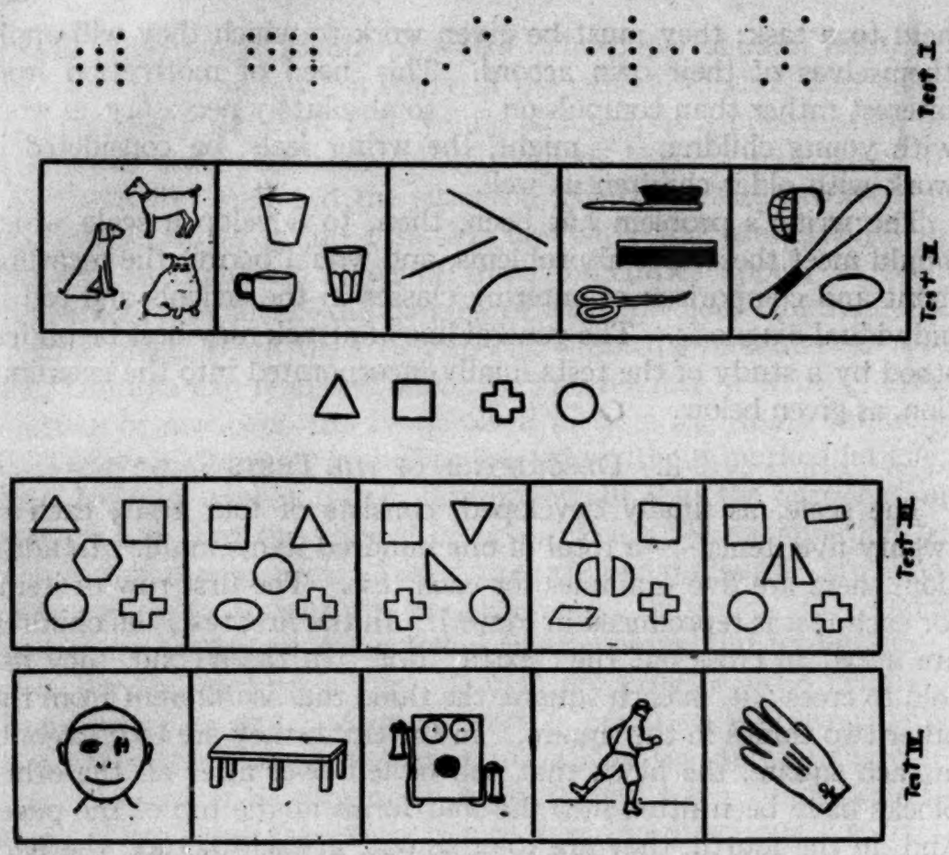
held to a task; they must be given work to which they will apply themselves of their own accord. This need of motivation from interest rather than compulsion – – so absolutely necessary in work with young children – – might, the writer feels, be considered in work with older children as well.

The writer's problem has been, then, to develop a scale which would meet these special problems, and would permit the measurement and comparison of entering classes in the schools and aid in individual diagnosis. The general line of attack may best be understood by a study of the tests finally incorporated into the examination, as given below.

2. Description of the Tests

The scale, as finally developed, consists of four tests, each of twenty-five items – – a total of one hundred items in all. In addition, there are five examples for each test. The first row of items for each test is reproduced in Plate I. In the first test, the children are asked to cross out the "extra" dot. In the second, they are told to cross out, in each square, the thing that is different from the other two things in the square. In the third, they are to cross out, in each square, the block that will be left over after all the other blocks have been fitted into the four forms at the top of the page. And, in the fourth, they are to cross out, in each picture, the part that is wrong. The examination appears on a four page folder with one test on each page, the pages being 11" by 8½."

(1) The studied avoidance of literacy in these tests is evident; two only of the hundred items involve any school acquirement and, in these two cases, the problem presented is not one wholly dependent upon direct teaching. In this respect, the writer feels that the examination is even more free from information gained from schooling than the Binet tests for the same ages. (2) Fatigue is kept negligible by calling for only one very simple response from the child. In each test he merely crosses out – – using any kind of a mark – – a certain object, part or dot. No fine eye-hand coordinations (so difficult for young children to make) are demanded, as any mark on approximately the right spot, will answer the purpose. By basing all four tests on this same "cross-out" principle the clerical labor of the child in taking the examination, is practically nil. Moreover, confusion from test to test is avoided. As a result of all these factors, it is possible to present the entire examination of 100 clear cut problems in less than twenty-five minutes; the



children actually work at each test only three minutes. (3) The children take an active part in the directions – – in fact, their answers constitute the *real* directions, as the majority of the children pay little attention to the final remarks of the examiner. The question-answer method is used. It might seem, in these days of absolute standardization of directions, that such directions would be loose and not capable of standardization. But, by using questions, the answers to which are very simple, the directions are kept practically standard. The examiner asks the same questions of every grade and from every grade, from the kindergarten thru the fourth, she receives the same, practically invariable, answers from the children. This method keeps the examiner in touch with the children, and has the further merit of automatically regulating the speed of giving the directions according to the time demanded by the class in understanding them. The directions for Test I are given verbatim below.

Look at the first group of dots at the top of the page. (Hold up copy and indicate general position). Can anyone see an extra dot – – one that is out of place? . . . Where is it? . . . Yes, above all the others. Now I want you to take your pencils and cross out that extra dot. Just draw a line through it. (Make sure that *all* the children get this first example correctly crossed out).

Now look at the next group of dots, just *beside* the first group. Where is the extra dot?... Yes, below all the others; everyone cross out the extra dot. Draw a line through it.

Now look at the next group, just beside the last one. Where is the extra dot?... Yes, above the line. Now all cross it out.

Now look at the next group. Where is the extra dot?... Yes, the last one on this side (gesture to the children's right). Cross it out.

Now look at the last group. Where is the extra dot?... Yes, the one just below the middle dot. Cross it out.

Now, everyone attention! Finish the other groups of dots on the page in the same way. In each group there is just *one* extra dot. Cross it out. (Time allowed pupils – – 3 minutes).

(4) The scale is carefully integrated. For each test, the five examples are in the same place on the page and are taken up in the same order; for each test, the directions develop these examples in the same way; for each test, the items are arranged the same way on the page. Above all, in each test, the child responds in the same way, – – he crosses something out. In fact, the most important element in the integration of the four tests is the use of this "cross-out" principle [2] in all the tests. (5) Most important of all is the development and sustaining of the childrens' interest throughout the examination. In the first place, the materials of the tests are pictures and blocks, in which children have a natural interest. The directions are given in such a way as to stimulate interest and rivalry among the children; the examples are interesting and the answers to the questions concerning them are so simple that every child can volunteer, thus having the pleasure of feeling he is contributing something to the success of the explanation. From the child's point of view, he and the examiner seem to be playing a game, in which he joins because he wishes to do so. The writer feels that there is a subtle but fundamental difference between an examination thus motivated by interest and the more formal type of examination usual in mental measurements.

3. Development of the Scale

Work on the scale was begun in the summer of 1918. A trial of various forms of tests was made with a mixed class of second and third grade children in the observation classes of the summer school of Indiana University. This class was composed of a few very bright children who were skipping a half grade and a few very dull children who needed extra training to pass the grade they had

[2] For a more complete discussion of the usefulness of this principle in test building, see Pressey, S. L. and L. W. *Cross-Out Tests with Suggestions as to a Group Scale of the Emotions.* J. of Appl. Psychol. June, 1919.

been in; it therefore presented an excellent group for experimental purposes. A tentative selection of tests was made from the results with these children.[3]

Experimentation during the fall was much hindered by the influenza epidemic, but some 200 children in country schools were tested and more systematic work was done in the primary grades of a small town. This experimentation served well to define the problem, make trial of methods and to lead to a more definite selection of the individual tests.

In the experiments thus far, the results had been obtained with pictures and forms drawn on the mimeoscope. The four tests that had been decided upon were now made into plates and blanks were printed for systematic study. In this form, the scale was given to all the children in the first, second and third grades of City A – – a city of about 12,000; 475 children were tested in all. The results were carefully worked over, the relation of the test findings to age-grade status, teachers estimates, Binet ratings and other criteria of ability studied, and the results of each test tabulated by items. On the basis of these findings a considerable number of new items were introduced (some entirely new, some taken from the earlier mimeographed sheets) to make the scale more evenly

[3] A total of some fifteen tests was tried at one time or another. Certain of these, though unsatisfactory for the writer's problem, might be of use to other workers and deserve brief mention. One (1) was a letter-pattern test in which the children were supposed to fill in the missing letters:—

L	M	C	B	A	B
L	M		A	B	A
L	M	C	B		B

The test was discarded because it demanded both knowledge of the alphabet and ability to write the letters. It should have possibilities, however, with older childen. Another test (2) involved discrimination of geometrical forms; rows of five forms, such as a square, a rectangle, a paralleogram, a rhomboid and a triangle, were presented. The children were asked to cross out, in each line, the form that was different from the other four. The test was rejected as being little more than a test of sensory discrimination. A third (3) test, of some interest was a directions test with pictures. The first item showed a face and the examiner directed that the children cross out the left eye; another picture showed a house and the directions were to cross out the window on the second floor that was not over a door or another window. The test was rejected because the children were so unused to disciplinary control that they would persist in looking around at their neighbors' papers between directions. A fourth test, (4) of some value for other purposes, was a substitution test. The key consisted of four geometrical figures, (a square, a cross, a triangle and a circle), with the silhouette of a lion in the first form, an elephant in the second, a kangaroo in the third and an eagle in the fourth. Below, on the page, were rows of the same forms, five in each row, with the animals again in the forms, but with one animal in each row in a different form from that in which he appeared in the key. The children were told to cross out, in each row, the animal that was in "the wrong cage." The children showed great interest in the test, but it was discarded because it was essentially a rate test and exact timing with young children is impossible.

graded in difficulty. The scale in this revised form was next given to all the children in the first three grades of City B; – – a total of 431 children.[4] Results from this trial were also tabulated by test and by item.

It was desired that the beginning of each test should be easy enough for any first grade child, that the end should be hard enough for use in the fourth grade if necessary, and that, between these two limits, the items should be arranged in order of difficulty, the steps between the successive items being approximately equal. After the analysis by item of the papers from City B was made, the items were graded on a probable error scale, using the second grade as a basis, and were then arranged in order of hardness. This treatment of the results showed that the steps were somewhat uneven and that there were some duplicate items (*i.e.* items of the same hardness). Other items, whose difficulty was already known from previous experimentation, were put in to help even the steps. The final product was a scale, the four tests of which were quite evenly graded in difficulty.

The results of this last revision were given to all the children in the first three grades of City C, – – with a total of 402 cases. The results of the survey with this final form are given below [5]

4. Results: First Norms and Validation

The age and grade distributions, with medians and percentiles, by total score, are given in Table I. The grade medians for the separate tests are given in Table II. These norms should be considered tentative as they are based on relatively few cases; but they will serve as a guide until further experimentation has been done and more cases tested.

---

[4] As indication of the ease and rapidity with which the examination can be given and scored, it might be said that all of these 431 children were examined by the writer and her husband in one day, between 8 a. m. and 3 p. m. and that all the blanks were scored by the writer and two students before 6 o'clock of the same day.

[5] Four items have since been changed because of difficulties in scoring. It is calculated that this slight change in items will raise the norms for the first grade about a point and between one and two points for second and third grades.

5-a. An intensive statistical study of the reliability and validity of the scale and separate tests is now being made. The data given below bearing upon the validity of the scale should be considered only a preliminary scale.

TABLE I

Primer Scale—Distribution of Scores

Children in grades 1–3 inclusive (402 cases) City C.

	I		II		III		Age				
Score	B	A	B	A	B	A	6	7	8	9	Score
95–99											95–99
90–94											90–94
85–89					1	2		1	1	1	85–89
80–84					4	5			4	2	80–84
75–79				2	5	16			9	8	75–79
70–74		2	3	4	5	17		7	7	12	70–74
65–59		2	8	16	12	9	1	15	16	10	65–69
60–64	2	8	4	13	13	14	5	14	12	15	60–64
55–59	2	4	5	8	5	4	6	7	7	5	55–59
50–54	7	9	12	13	4	4	6	22	13	5	50–54
45–49	5	17	8	6	2	5	13	14	9	5	45–59
40–44	5	11	4	2	4		10	5	6	2	40–44
35–39	5	5	7	2			8	7	1	2	35–39
30–34	5	5	2			2	8	4	3		30–34
25–29	3	6	1		2		6	2	3		25–29
20–24	2	2			1		2	2	1	1	20–24
15–19	7	3					6		3		15–19
10–14	6						5	1			10–14
5–9	6	3					6				5–9
0—4	11	1	1		1		9	1	1	1	0—4
Number	66	78	55	66	59	78	91	102	96	69	
	144		121		137						
Median	26.7	45.9	50.9	60.8	64.0	70.3	37.2	63.5	61.0		
25%	9.6	34.5	43.5	52.4	55.8	61.4	17.2	46.3	48.7		
75%	49.5	54.4	61.5	61.7	70.2	76.1	48.2	64.4	69.8		

Note: It should be noted that the curves for the IB and 6 year distributions are not normal because of the number of zero scores – – due to not understanding the directions. In calculating the medians for 7 and 8 years, the one seven year old and four eight year olds above the third grade, (and therefore not tested) have been counted in as being above the median. The other percentiles were calculated from the figures presented.

TABLE II.

Medians for each test.

Grade	No.	Test I	Test II	Test III	Test IV
1	144	10.3	10.5	9.2	8.9
2	121	13.3	15.5	15.8	14.7
3	137	15.7	17.0	20.3	16.1

In this last city all data possible were obtained bearing on the question of the validity of the scale. Light regarding the validity was gained from five sources: (1) correlation with teachers' estimates, (2) correlation with Binet ratings, (3) location on the total distribution of the scores made by children having a Binet I. Q. of 123 or more and by those having an I. Q. of 76 or less, (4) comparison of "slow" and "fast" sections, and (5) comparison of the entering classes in the different schools.

Seventeen correlations were obtained between the test scores and teachers rankings of pupils according to ability; these rankings were obtained from all teachers in City C, except from some few substitutes who did not know the children well. The average correlation was .47.

Three correlations were possible between the test scores and mental age on the Stanford Revision. For 64 unselected six year old children, the correlation is .64; for a group of 33 unselected seven year olds, it was .69; and for a similarly unselected group of 29 eight year olds, it was .79. [6] The product-moments method was used throughout. The writer feels that a higher correlation is hardly to be expected between a group test and an individual test -- at least in work with young children -- because the test situation is so different in the two methods. Bashful children do better when submerged in a group and lazy children do better when constantly prodded by the Binet worker, etc.

Among the 6, 7 and 8 year old children who were given Binet examinations in two of the three cities tested, 24 were found whose I. Q. was over 123. Twenty-two of these make group test scores in the highest ten per cent. for their age. The other two make very low scores that are obviously unreliable. Considering the difficulty of testing such young children at all, the scale seems reasonably reliable at the upper end of the distribution. There were only eleven children with an I. Q. of 76 or below; all of these made group scores in the lowest ten per cent. for their age. Since the two extremes of the distributions seem reliable, the scale would appear to be fairly satisfactory as an aid in individual diagnosis, as the cases to be further studied are usually located in the extremes.

Validation of a different sort, but of perhaps more direct implication, appears in the comparison of certain groups, such as sections and classes. In many of the classes tested, the teacher had already divided her class into "slow" and "fast" sections on the basis of her estimate of their ability. It is not to be expected that her judgment would be absolutely accurate; but the fact that a child has, throughout the year, been able to hold his place in a "fast" section when he might at any time have been transferred to a slower one, is surely of some significance. On the average, 76 per cent. of the children in the "fast" sections score above the median of the

[6] The ages were kept separate since it is obvious (in spite of frequent practice to the contrary) that combining several ages adds a spurious element to the correlation. Mental age was used instead of I. Q. since, for such a correlation, absolute score on both tests is necessary. No per cent. rating with the group rating was possible.

corresponding "slow" section and, in all classes but one, 100 per cent. of the children in the "fast" sections score above the twenty-five percentile for the corresponding "slow" sections.

Similar evidence appears when the children of a school whose pupils are known to be distinctly below average are compared with the children of another school that is known to draw from a superior neighborhood. Such evidence has been found in all three of the cities tested. In City C there are three ward schools: one of these is situated in the best residential district, another in a down-town district, and the third in the poorest part of the city. It had been felt for some years that the children attending the three schools were very different in their mental capacities – – that the "pupil material" of the three schools was very unlike. The results in Table III show the per cent. of children in the second and third schools who score above the median for the corresponding grade and age in the best school – – *i. e.* the one situated in the best residential district. If it can be assumed that the school officials were

Grade.	No.	Sch. 2.	No.	Sch. 3.	Age.	No.	Sch. 2.	No.	Sch. 3.
1	37	44%	48	40%	6 and 7	54	44%	57	42

right, the scale seems to have measured the differences already felt. It seems evident that the "pupil material" of the three schools is really different. In City B, there is a ward school that is attended almost exclusively by the children of university teachers and other professional men; another school in the same city is attended entirely by the children of factory and quarry workers. Only 24 per cent. of the children in the second school tested above the age medians of the children attending the first school; there was only one child in the second school who tested above the seventy-five percentile for children of his age in the first school. The second school has been considered for years as inferior to the first and, as was the case in City C, the scale has found and measured the difference. Similar facts can be presented for City A, in which 74 per cent of the children attending the "best" school score above their corresponding age medians for the children in the "worst" school, – – which in this case, draws from a neighborhood of the cheapest kind of quarry labor.

Such evidence as has been presented in the last two paragraphs is crude enough; but it is direct evidence and it is also an evidence of the usefulness of the scale in dealing with actual educational

problems. It shows that the scale is differentiating, to some extent at least, the ability that makes a pupil, or a group of pupils, better or less able than the average to do successful school work.

5. Discussion

The scale described in the previous pages is, as far as the writer knows, the first attempt at a group scale of intelligence for use with first grade children. It is, therefore, pioneer work. The writer feels that some discussion is needed concerning (1) certain special reasons why tests of intelligence for use with first grade children should be of unusual value and (2) certain special problems which appear in test work with young children and which need further study.

(1) Any school makes more or less adjustment to the mentality of the pupils attending it. That these adjustments may be as thorough as possible it would seem necessary to obtain some measure of the intelligence of this "pupil material." The only grade in which this measurement can be satisfactorily made is the first. In the upper grades some adjustment and considerable selection has already taken place; in fact, measurements of the upper grades may be little more than measurements of the extent of such adjustment and selection.[7] Comparison between schools on the basis of such results is impossible because of the different amounts of selection in the different schools. But that such differences are marked and important the writer has especially tried to show. Surely, measurement of the distribution of abilities obtained from the upper grades only gives a very partial and incomplete picture of the total situation in a school or school system. Scientific study of individual differences should be begun as soon as possible after the child's entrance.

(2) The writer feels that there are certain very special problems in building tests for young children. The most fundamental of these is the control of interest, but of nearly equal importance is the avoidance of confusion from test to test. Interest is controlled principally through the directions. There might be some doubt as to whether or not the question-answer method is the best way of doing it. This method keeps the examiner in touch with the class, keeps the class in touch with her and seems to work well. But other methods should certainly be tried. The writer is certain only of this, that in work with very young children a special problem is presented as regards directions. With older children it is enough

[7] Pressey, S. L. *A Comparison of Two Cities by Means of a Group Scale of Intelligence.* Educ. Ad. and Super. Vol. 5. pp. 53-62, 1919.

that the directions be clear; with young children there is the further fundamental requirement that there shall be control of interest. Any other methods used must recognize this fact.

Confusion from test to test has been avoided by close integration of the tests. Such unification greatly facilitates the work. It may be that, in thus simplifying the procedure, certain elements of difficulty which should appear in such an examination have been left out of the tests; the ability to shift from one test to the next may be one of the best indications of intelligence! But, at least, such unity is worth trying. It has seemed to the writer that it tended to a greater concentration of effort on the real problem of the tests, with less attention to irrelevant matters of test form.

6. Summary

The paper may be briefly summarized:

1. It is urged that a group scale of intelligence is necessary in the primary grades in order (a) to permit comparison of the "pupil material" entering different schools and (b) to aid in individual diagnosis.

2. It was found that, in building a scale for young children, certain special requirements must be fulfilled: (a) all elements of school training must be avoided, (b) fatigue must be reduced to a minimum, (c) the directions must be such that the children can take an active part in them, (d) the tests of the scale must be carefully integrated, and (e) the work must be motivated and maintained by interest rather than by disciplinary control.

3. A brief survey scale is presented which endeavors to meet these requirements. The scale consists of four tests, each of twenty-five items; the items are arranged in order of difficulty on a Probable Error scale.

4. Tentative norms are given, and evidence presented to show the validity of the scale as a measure of general ability.

A CLASSIFIED SCALE FOR MEASURING INTELLIGENCE

CARLETON W. WASHBURNE
San Francisco State Normal School

The rapidly increasing use of intelligence testing in schools necessitates a modification of existing tests. All tests now used measure intelligence more or less efficiently, but fail to diagnose in any consistent way the specific lacks or excellences of any individual intelligence.

School psychologists, to be worth their hire, must be able to say not merely, "This child is superior, this one inferior," but to say specifically, for instance, "This child is normal in most mental functions not involving immediate word memory, but he is deficient in that. Do not expect it in his work—build on the normal functions and see if you can train this deficient one."

All of us do, of course, make an effort at this sort of diagnosis, but existing scales for measuring intelligence are not adapted to our uses. For this reason I have undertaken to make a practical diagnostic scale which will not only measure the general intelligence but will also indicate in what particular functions a child is above or below normal.

The Problem and Plan of Attack

The problem in constructing the classified scale has been:

a. To make the test specifically diagnostic as well as general.
b. To make a classification which would be practical rather than technically psychological, and which would, therefore, be suggestive to teachers.
c. To have the results of the test such that they could be compared with the results of other widely used intelligence tests.
d. To use the results of the extensive investigations and revisions which have already been made by others.
e. To keep the test within the time and skill limits already set by successful intelligence measurement scales.

It is patent that the age-classification used so commonly is general and not diagnostic. The Binet scale was not planned for specific diagnosis. It set out deliberately to measure *general* intelligence. A scale should be devised which would cover the most significant

of the mental functions each in a thorough way. But as Wallin has pointed out, it is better not to complicate the field further by adding more revisions until a foundation or organization of psychologists undertakes the job and does it far more thoroughly than any one person or small group could possibly do it. Consequently I considered it better to use the best current revision of the Binet Scale as a basis for the classification, leaving the procedure, materials and method of scoring intact and simply rearranging the tests according to mental function rather than according to age. My classification is frankly a make-shift until such time as a new general revision can be made which will include adequate tests diagnostic of each function at different age levels.

Meanwhile, the revised Binet scale is fairly susceptible to a diagnostic reclassification. Yerkes and his co-workers, in their point scale, attempted one such reclassification. Their groupings, however, are distinctly technical. To the ordinary teacher and even to the trained psychologist they are not especially enlightening when it comes to giving practical suggestions as to the education of the pupil tested. Furthermore, the Yerkes scale came out just too soon to take advantage of the thorough work and revision carried on by Terman.

Terman's Stanford Revision of the Binet-Simon scale, in spite of several obvious faults, certainly is the best available general intelligence test for school use, and has an easily available, completely worked out procedure and scoring.

For this reason I have chosen the Stanford Revision as the basis for my diagnostic classification, and have so arranged my scale as to make use of Terman's procedure and scoring directions and to secure an I. Q. which is identical in each case with that which would be secured by the straight Stanford Revision.

The method of making the classification was necessarily rough. None of us knows exactly what mental functions are used in the passing of any one test. Different individuals doubtless vary somewhat among themselves in this particular, and quite possibly every mental act is conditioned by all mental functions. Nevertheless, different functions clearly predominate in different acts. My effort, therefore, has been to group together under the head of one function all tests the passing of which clearly involved the use of that function; to include a test in more than one group when more than one function was apparently of importance in passing it; and to make

the groupings not from the standpoint of technical psychology but rather from that of practical teaching problems.

In order to group the tests properly each test in the Stanford Revision of the Binet-Simon Scale was carefully examined. Terman's remarks on each test were considered and compared with the observations I had been making with this grouping in view. A temporary classification was made on this basis, then after about a hundred children had been tested by this preliminary classified scale, the results were studied and the classification thoroughly revised. The result was the following classified scale:

THE CLASSIFIED SCALE

(Note: Hyphonated numbers refer to Terman's age groupings, e. g., 3-1 means that the test so numbered is test no. 1 of year 3, of the Stanford Scale; 8–A1 is the first alternative test in the 8-year group, etc.)

No.	*Title of test*	*Grouping*	*Cross reference*
		I. NAMING THINGS	
1.	3–1	Points to parts of body	
2.	3–2	Names familiar objects	
3.	3–3	Enumeration of objects in picture	
4.	(3–4)	Gives sex	See II, 1
5.	3–5	Gives last name	
6.	5–2	Names colors	
7.	6–1	Right and left	Also II, 3
8.	6–5	Names three of four coins	Also II, 4
9.	8–A1	Names six coins	Also II, 9
		II. COMMON KNOWLEDGE	
1.	3–4	Gives sex	Also I, 4
2.	5–A1	Gives age	
3.	(6–1)	Right and left	See I, 7
4.	(6–5)	Names three of four coins	See I, 8
5.	6–A1	Morning or afternoon	
6.	(7–1)	Number of fingers	See XIII, 3
7.	(7–4)	Ties bowknot	See XIX, 1
8.	7–A1	Names days of week	
9.	(8–A1)	Names six coins	See I, 9
10.	(8–A2)	Writes to dictation	See XIX, 2
11.	9–1	Gives date	
12.	9–A1	Names months of year	
13.	(10–4)	Reading and report	See XIX, 3
14.	(12–2)	Meaning of abstract words	See XVI, 3
15.	14–3	Differences between president and king	Also VII, 8
16.	(14–4)	Problems of fact	See IX, 7
17.	(16–3)	Differences in abstract words	See XVI, 4
18.	(18–4)	Repeats thought of passage	See XV, 17

III. IMMEDIATE AUDITORY MEMORY—DIGITS

1. 3–A1 Repeats three digits
2. 4–6 Repeats four digits
3. 7–3 Repeats five digits
4. 10–A1 Repeats six digits
5. 14–A1 Repeats seven digits
6. 18–3 Repeats eight digits

IV. IMMEDIATE AUDITORY MEMORY—SENTENCES

1. 3–6 Repeats 6-7 syllables
2. 4–A1 Repeats 12-13 syllables
3. (5–6) Three commissions.....See XV, 4
4. 6–6 Repeats 16–18 syllables
5. 10–A2 Repeats 20-22 syllables
6. 16–A1 Repeats 28 syllables

V. VOCABULARY

1. 8–6 20 words
2. 10–1 30 words
3. 12–1 40 words
4. 14–1 50 words
5. 16–1 65 words
6. 18–1 75 words

VI. VISUAL MEMORY (PERCEPTION)

1. (6–2) Mutilated pictures.....(See IX, 1)
2. 10–3 Copies designs from memory.....Also XVIII, 6
3. (10–4) Reading and report.....See XIX, 3
4. (12–4) Dissected sentences.....See X,10
5. 14–6 Reversing hands of clock.....Also IX, 8
6. 16–6 Code.....Also X, 12
7. (18–2) Paper cutting.....See IX, 10

VII. QUALITATIVE COMPARISON

1. 5–3 Aesthetic comparison
2. (5–5) Divided rectangle.....See VIII, 3
3. (6–2) Mutilated pictures.....See IX, 1
4. 7–5 Gives differences
5. (8–4) Gives similarities.....See XVII, 1
6. (9–6) Rhymes.....See X, 6
7. (12–8) Gives similarities, three things.....See XVII, 4
8. (14–3) Gives differences between president and king.....See II, 15
9. (16–3) Gives differences between abstract words.....See XVI, 4

VIII. DISCRIMINATION OF FORM

1. 4–2 Discrimination of forms
2. (4–4) Copies square.....See XVIII, 1
3. 5–5 Divided rectangle.....Also VII, 2; X, 1; XI, 1; XVIII,2
4. (7–6) Copies diamond.....See XVIII, 4

IX. IMAGINATION (PERCEPTION)

1. 6–2 Mutilated pictures.....Also VI, 1; VII, 3; X, 2
2. 7–2 Pictures, description
3. (8–1) Ball and field, inferior plan.....See X, 3
4. (9–5) Three words in a sentence.....See X, 5
5. (12–3) Ball and field, superior plan.....See X, 9
6. 12–7 Pictures, interpretation
7. 14–4 Problems of fact.....Also, II, 16; X, 11
8. (14–6) Reversing hands of clock.....See VI, 5
9. (16–4) Enclosed boxes.....See XIV, 5
10. 18–2 Binet's paper cutting test.....Also VI, 7
11. (18–6) Ingenuity test.....See X, 13

X. INVENTIVENESS AND INGENUITY

1. (5–5) Divided rectangle.....See VIII, 3
2. (6–2) Mutilated pictures.....See IX, 1
3. 8–1 Ball and field, inferior plan.....Also IX, 3; XV, 6
4. (9–2) Arranges weights.....See XII, 3
5. 9–5 Three words in sentence.....Also IX, 4; XV, 8
6. 9–6 Rhymes.....Also VII, 6
7. 10–6 Sixty words
8. 10–A3 Form board.....Also XI, 2
9. 12–3 Ball and field, superior plan.....Also IX, 5; XV, 12
10. 12–4 Dissected sentences.....Also VI, 4; XV, 13
11. (14–4) Problems of fact.....See IX, 7
12. (16–6) Code.....See VI, 6
13. 18–6 Ingenuity test.....Also IX, 11; XIV, 6

XI. PATIENCE OR PERSISTENCE

1. (5–5) Divided rectangle.....See VIII, 3
2. (10–A3) Form board.....See X, 8

XII. QUANTITATIVE COMPARISON

1. (4–1) Compares lines.....See XV, 1
2. 5–1 Compares two weights.....Also XV, 3
3. 9–2 Arranges five weights.....Also X, 4

XIII. USING AND MANIPULATING NUMBERS

1. 4–3 Counts four pennies
2. 6–3 Counts thirteen pnniees
3. 7–1 Number of fingers.....Also II, 6
4. 7–A2 Repeats three digits backward
5. 8–2 Counts backward, 20–1
6. 9–4 Repeats four digits backward
7. 12–6 Repeats five digits backward
8. 16–5 Repeats six digits backward
9. 18–5 Repeats seven digits backward

XIV. ARITHMETICAL REASONING (ALSO XIII)

1. 9–3 Makes change
2. 9–A2 Values of stamps
3. (14–2) Induction test.....See XVII, 5
4. 14–5 Arithmetical reasoning
5. 16–4 Problem of enclosed boxes.....Also IX,9; XVII,8
6. (18–6) Ingenuity test.....See X, 13

XV. COMPREHENSION (ALSO IMAGINATION)

1. 4–1 Compares lines.....Also XII, 1
2. 4–5 Comprehension, first degree
3. (5–1) Compares two weights.....See XII, 2
4. 5–6 Three commissions.....Also IV, 3
5. 6–4 Comprehension, second degree
6. (8–1) Ball and field, inferior plan.....See X, 3
7. 8–3 Comprehension, third degree
8. (9–5) Three words in sentence.....See X, 5
9. 10–2 Detects absurdities
10. (10–4) Reading and report.....See XIX, 3
11. 10–5 Comprehension, fourth degree
12. (12–3) Ball and field, superior plan.....See X, 9
13. (12–4) Dissected sentences.....See X, 10
14. (12–5) Interpretation of fables (Score 4).....See XVII, 3
15. (16–2) Interpretation of fables (Score 8).....See XVII, 6
16. 16–A2 Comprehension of physical relations
17. 18–4 Repeats thought of passage heard.....Also II,18

XVI. DEFINING (ASSOCIATION AND ANALYSIS)

1. 5–4 Definitions, use or better
2. 8–5 Definitions, superior to use
3. 12–2 Meaning of abstract words.....Also II, 14; XVII, 2
4. 16–3 Differences between abstract words.....Also II, 17; VII, 9; XVII, 7

XVII. ABSTRACTING COMMON ELEMENT–GENERALIZATION; INTERPRETATION

1. 8–4 Gives similarities, two things.....Also VII, 5
2. (12–2) Meaning of abstract words.....See XVI, 3
3. 12–5 Interpretation of fables, (Score 4).....Also XV,14
4. 12–8 Gives similarities, three things.....Also VII, 7
5. 14–2 Induction test.....Also XIV, 3
6. 16–2 Interpretation of fables, (Score 8).....Also XV, 15
7. (16–3) Differences between abstract words.....See XVI, 4
8. (14–4) Problem of enclosed boxes.....See XIV, 5

XVIII. MOTOR COORDINATION

1. 4–4 Copies square.....Also VIII, 2
2. (5–5) Divided rectangle.....See VIII, 3
3. (7–4) Ties bowknot.....See XIX, 1
4. 7–6 Copies diamond.....Also VIII, 4
5. (8–A2) Writes to dictation.....See XIX, 2
6. (10–3) Copies designs from memory.....See VI, 2

XIX. SPECIA TRAINING

1. 7–4 Ties bowknot.....Also II, 7; XVIII, 3
2. 8–A1 Writes to dictation.....Also II, 10; XVIII,5
3. 10–4 Reading and report.....Also II, 13; VI, 3; XV, 10

How Tests are Scored

As can readily be seen above, most of the tests fell under two or more heads. But in calculating the intelligence quotient the same test could not be scored twice, of course; neither could alternative tests be scored except where a regular test was omitted. To provide for this a simple device was used—two columns for scoring were made, one a months-credit column in which to record the number of months to be credited in calculating I. Q., and the other the function-age column in which to note what age level the pupil reaches in each function. In the function-age column there is a blank for score opposite each test every time it appears; but in the months-credit column there is only a blank for scoring each test once and no blank for scoring alternative tests. Thus the blank for scoring reading and report (XIX, 3 above) would occur but once in the months-credit column (under Special Training, XIX) but would appear four times in the function-age column, i. e., each time the test appeared in the scale (under Common Knowledge, II; under Visual Memory, VI; and under Comprehension, XV as well as under Special Training, XIX). This same device makes it easy to tell at a glance whether each test has been given once, for an unfilled blank in the months-credit column will immediately indicate that the test opposite has not yet been given. Naturally we only give each test once, the additional scoring in the function age column being done after the test has been completed.

Perhaps a sample page from the record booklet will make this easier to visualize:

Sample Page from Record Booklet

IV. IMMEDIATE AUDITORY MEMORY (SENTENCES)

Mo. Cr.	Fnct-Age	
1)...*	...3-6.	Repeats 6-7 syllables (1 of 3) a. I have a little dog..... b. The dog runs after the cat..... c. In summer the sun is hot.....
2)	..,4-A1.	Repeats 12-13 syllables (1 of 3 absolutely correct or 2 of 3 one error each) a. The boy's name is John. He is a very good boy. b. When the train passes you will hear the whistle blow. c. We are going to have a good time in the country.
3)	...(5-6).	Three commissions. (See XV, 4)
4)...	...6-6	Repeats 16-18 syllables. Instructions like 2) above. a. We are having a fine time. We found a little mouse in the trap. b. Walter had a fine time on his vacation. He went fishing every day. c. We will go out for a long walk. Please give me my pretty straw hat.
5)	...10-A2	Repeats 20-22 syllables. Instructions as above. a. The apple tree makes a cool pleasant shade on the ground where the children are playing. b. It is nearly half-past one o'clock; the house is very quiet and the cat has gone to sleep. c. In summer the days are very warm and fine; in winter it snows and I am cold.
6)	...16-A1.	Repeats 28 syllables. (1 of 2 absolutely correct. (5 months) a. Walter likes very much to go on visits to his grandmother, because she always tells him many funny stories. b. Yesterday I saw a pretty little dog in the street. It had curly brown hair, short legs, and a long tail.

V. VOCABULARY

1)...*	... 8-6.	20 words...
2)...*	...10-1.	30 words...
3)...*†	...12-1.	40 words...(3 or 4 months)
4)...*†	...14-1.	50 words...(4 or 6 months)
5)...*†	...16-1.	65 words...(5 or 7.5 months)
6)...*†	...18-1.	75 words...(6 or 9 months)

VI. VISUAL MEMORY (PERCEPTION)

1)...	...(6-2)	Mutilated pictures (See IX, 1)
2)...	...10-3.	Copies designs from memory. (1 correct, other half correct.) Expose 10 seconds. (Also XVIII, 6) a...b....
3)...	...(10-4)	Reading and report (See XIX, 3)
4)...	...(12-4)	Dissected sentences (See X, 10)
5)...†	...14-6.	Reversing hands of clock (2-3. Error not over 3 or 4 min.) (Also IX, 8) (*4 months*) 6.22... Time req.... 8.10... Time req.... 2.46... Time req....
6)...†	...16-6	Code. Writes "Come quickly," (2 errors, six minutes. Omission of dot counts half error. Illustrate with "war" and "spy") (Also X,12) (*5 months*) Errors...C. O. M. E. Q. U. I. C. K. L. Y. Time...
7)	(18-2)	...Paper cutting (See IX, 10)

In the left-hand (months-credit) column of the booklet the number of months to be credited for each test is recorded whenever there is a blank line in this column opposite the test passed. The number of months to each test is the number assigned by Terman in his Stanford Revision—*i. e.*, two months for each test below the twelve-year level, three months for tests of the twelve-year level, four months for the fourteen-year level, etc. For the convenience of the examiner a mark † is placed by the months-credit blank opposite each test that is to receive more than two months credit, and at the end of each such test the number of months to be credited is shown in parenthesis, thus: (*4 months*) as in test VI, 5 above, or (*3 or 4 months*), as in test V, 3 above. In cases like the latter, where two numbers are given in the parenthesis, it is ordinarily the smaller of the two (*e. g.*, *3*) that is credited; but in case the examiner, to save time, is giving only those tests which are marked with an asterisk (*) he gives the larger of these two numbers (*e. g.*, 4) instead. In this last event—*i. e.*, when only the asterisk tests are give—all tests below the twelve-year level receive three months credit each instead of two.

In the right-hand (function-age) column, each test given is recorded either plus or minus according to whether the pupil passes it or fails to do so. The hyphenated numbers immediately preceding each test (*e. g.*, *3*-6), as above noted, indicate the age level and Terman's test number—thus 3-6 means test No. 6 (Stanford Scale) of the 3-year group. So the first of the two hyphenated numbers always shows the approximate age level (*e. g.*, three years) at which most children are able to pass the test opposite. The highest age level reached by a child in any function, therefore, indicates roughly his mental age in regard to the particular function under which the test is listed. This we call his "function-age" for that function.

Let us suppose, for instance, that we are testing a nine-year-old child. We come to the page in the record booklet just reproduced above. There is no test in Immediate Auditory Memory, Sentences, for a nine-year child. We try the ten-year test (no. 5) therefore. (For getting function-age the alternative tests, like this one, are often helpful, even though they play no part in determining I. Q.) He fails on it. We place a minus sign on the line in the function-age column opposite and pass back to the six-year test (no. 4). He passes it, but just barely. We note this and give him plus in the

function-age column. In the months-credit column we credit him with two months for this test (no. 4) and also with two months for test no. 1 of the same group—for his passing the six-year test in this function is clear evidence that he could pass the three-year test in the same function. It is not so clear that he can pass the five-year test (no. 3) Three Commissions) however, since he just barely passed the six-year test, and expecially as no. 3 involves other functions. Consequently we would either give no. 3 now or remember to do so when we came to it later on—the note "See XV,2" shows that we will find the complete test as no. 2 under group XV of the scale.

Passing down to the next group we give the vocabulary test.* Suppose the child we are testing makes a score of 45 words. This number would be recorded, he would be marked plus in the function-age column opposite the twelve-year test, no. 3, and would be given three months credit for this test in the months-credit column and also two months each for nos. 1 and 2. For here again, the passing of the higher test in the function is evidence of the ability to pass the lower ones, this time rigorously so. The score for the child for this grouping is therefore seven months credit, and his function-age is thirteen years (for his score of 45 words is midway between that demanded for the twelve-year level and that demanded by the fourteen-year level.

Next, we pass on down to group VI, Visual Memory. Here as in Group IV there is a big gap between the six-and ten-year levels. Following our former policy, we try him first on the test nearest his physical age of nine years. This is no. 2, Copying designs from memory. Suppose he passes it. We credit him in both columns and go on down the scale. The next test is Reading and report, but as the hyphenated numbers are in parenthesis and the note after the test says "see XIX, 3" we leave test no. 3 until we come to group XIX. Similarly we skip, for the time being, no. 4, Dissected sentences, and try no. 5, Reversing the hands of the clock. Suppose he passes this, we record it in both columns and go on to no. 6, the Code test, on which he fails. He now has credit for six months

*The vocabulary tests could rather easily be grouped under common knowledge, and might even be related to auditory and visual memory. But since they form a well developed series in themselves, their significance is, I think, easier to grasp by themselves than mixed in with other tests.

(two for No. 2 and four for No. 5) in the months-credit column, and a function age of fourteen. Then let us suppose that when we come to group X (not shown on the sample page), he fails on the Dissected sentence test of twelve-year level. In scoring him afterwards this will make a minus in the function-age column opposite test 4 of group VI (Visual Memory) in contradiction to the plus in test 5 of the same group—in other words, although test 5 indicates that he has a function age of fourteen years in visual memory, test 4 indicates that it is not even twelve. Of course this is probably due to the fact that both tests are conditioned by other and different functions, or it may be due to an incorrect grouping. Anyway, we score his function-age as 14—in visual memory, the minus sign signifying that while he has passed a test of the fourteen-year level in this function, he nevertheless has failed a test of lower age level under the same function.

The scoring, then, of the child we have imagined testing, would be, so far as this sample page is concerned, as follows:

Group	*Function-age*	*months-credit*
IV Sentences	6	4
V Vocabulary	13	7
VI Visual memory	14—	6

At the end of the entire test all months' credits and function-ages are tabulated in this manner in blanks provided on the front cover (see sample below). The months' credits are added together (with 24 months extra for the first two years for which there are no tests), and this sum is the mental age of the one tested. It is, of course, identical with that obtained in the regular Stanford scale. I. Q. is calculated as usual by dividing this mental age by the child's physical age. The summary on the cover shows at a glance the mental age and intelligence quotient of the child, then, just as in the Terman booklet. But in addition to this it shows an analysis of the I. Q., giving a practical diagnosis of the subject's strong and weak points, indicating in what functions he is normal, in what ones deficient, and in what ones superior.

The Significance of the Scoring

It should be recognized clearly, however, that to say a child is six years old in immediate auditory memory of sentences and fourteen in visual memory is only an approximation to the truth, an

indication rather than a complete diagnosis. But it is an indication with considerable practical value, and taken into consideration with other indications may shed valuable light on the pupil's educational needs. Let me take for example a real case:

H. B. was a girl twelve years old doing fifth grade work poorly. At the request of her teachers I gave her an intelligence test. It yielded an I. Q. of 91—too near normal to justify nearly two years retardation, especially in a school where an individual system of instruction is in use. The I. Q. therefore gave no clew as to the difficulty. The classified scale, however, showed her function-ages to be as follows

SUMMARY OF RECORD OF H. B.

(From cover of record booklet) Age, 12–1; Mental age, 11.0; I. Q., 91.

No.	*Test group*	*Fnc. Age Mo. cr.*
I.	Naming things	+ 14
II.	Common knowledge	9 4
III.	Repeating digits	+ 4
IV.	Repeating sentences	+ 4
V.	Vocabulary	8 2
VI.	Visual memory	9? 0
VII.	Qualitative comparison	— 4
VIII.	Discrimination of form	+ 4
IX.	Imagination	12 7
X.	Inventiveness	12 11
XI.	Patience	+
XII.	Quantitative comparison	+ 6
XIII.	Manipulating numbers	+ 10
XIV.	Arithmetical reasoning	14— 2
XV.	Comprehension	12 12
XVI.	Defining	— 4
XVII.	Extracting common element	16— 14
XVIII.	Motor co-ordination	+ 4
XIX.	Special training	— 2
	First two years	24
	TOTAL MONTHS	132

Mental age 11 yrs. 0 months

Note: The plus sign is used wherever the gap in tests is such as to fail to give even an approximate function-age, but where no deficiency in this function is evident. The minus sign, not preceded by a number, indicates that while the gap in tests does not permit of an approximate function-age, there is an evident deficiency in this function.

Notes on H. B.: The functions in which she shows deficiency are common knowledge, vocabulary, defining, and special training (reading), all of which point strongly to poor environment, and visual memory and qualitative comparison where the lack of tests near her age make accurate scoring difficult—in each of these last two it was the failure in a single test which caused the apparent subnormality.

As just noted, H. B.'s lacks almost all pointed directly to poor home environment, while those functions which were largely innate, such as reasoning, extracting common element (induction), etc., were either normal or above normal. A further enquiry brought

out the fact that her father is an illiterate teamster, and that the girl's deficiency in common knowledge, vocabulary, etc., had led to marked diffidence and lack of self-confidence. A discussion of her case with her teachers gave them an immediate perception of the need for supplying the background of common knowledge, etc., which her home had failed to supply, and of bringing out her strong points in such a way as to restore her self-confidence.

The scoring in the function-age column admits of considerable latitude and discrimination, since it in no wise affects the I. Q. For example, in giving a bright 7-year old child the ball and field test, I noticed that as the child drew his path he made remarks somewhat as follows: "I'd go over here by the fence and look in the high grass, then I'd go over here and look in the bushes, then I'd look out in the middle," etc. The path was without plan and the test had to be scored minus in the months-credit column. But in the function-age column under imagination and under comprehension I gave the child full credit. Under ingenuity, however, I scored him minus in the function-age column as well as in the months-credit column. (The ball and field test is classed under all three functions, i. e., imagination, comprehension, and ingenuity). The difference between a complete failure and one which is only partial becomes, in this way, apparent on the score sheet, yet the rigorous determination of I. Q. is in no way vitiated.

The Results Achieved

For my own use, and that of my assistant, I have had mimeographed enough booklets to give the classified scale a satisfactory trial. A considerable number of children have now been tested by it, over a period of about six months covering a range from four years to fifteen (and a few adults) and in I. Q. from 73 to 165. The results of the tests have been very satisfactory—they have immediately suggested the type of work needed by the child and have helped materially in recommendations to teachers. The testing time, including the tabulation of results on the cover, ranges from twenty minutes for normal five-and six-year olds to an hour or more for very superior children of twelve and over. These times always include the full test, as I never use the short-cut test, indicated by asterisks, in my work. The test is as easily given as the regular Stanford Revision on which it is based. The scoring is only a little

more difficult and is very much more significant. The I. Q. yielded is of course identical with that yielded by the Stanford Revision—with this exception: The testing by the classified scale is likely to be more thorough and the I. Q. more accurate. For when there is a considerable gap in the Stanford Scale between two tests of the same function, as in the case of immediate auditory memory, for instance, a marked efficiency or deficiency in this function may never come to light. Such a child as we have hypothecated above, would perhaps pass all seven and eight year tests and therefore never be tested in the 6-year repetition of sentences in which he made such a poor showing. On the other hand, he might fail in all the twelve year tests and therefore never be tested in the 14-year test of visual memory, his marked superiority in this regard therefore never being indicated. In the classified scale the tendency is, so far as the tests will indicate this, to find the approximate age level of the child in each function.

Let me state once more that I frankly recognize that the grouping of the tests is only a rough approximation. With further testing it will doubtless have to be modified in various particulars. But I feel that there can be no question as to the advantages of testing intelligence by a classified scale rather than by the age-group scale heretofore used.

Conclusion

To sum up: (a) By classifying the tests of the Stanford Revision of the Binet-Simon scale according to the predominant function or functions used in passing them, it has been possible to make the scale specifically diagnostic as well as indicative of the general intelligence. (b) By making the groupings untechnical and practical the scale has been made suggestive to teachers. (c) By leaving the procedure and scoring of the Stanford Revision unchanged, the I. Q.'s yielded by the classified scale tests are strictly comparable with those obtained by the Stanford Revision, and (d) the results of Terman's extensive investigations have been used in their entirety. Finally (e) the test has proved itself to be within the same time and skill limits as those of the Stanford Revision.

The classified scale, therefore, solves in a practical way one important problem for the school psychologist—it shows him which particular mental functions are above or below normal and thereby gives one indication for the proper direction of educational effort.

A PRELIMINARY REPORT OF SOME GROUP TESTS OF GENERAL INTELLIGENCE

FRANCES LOWELL

University of Minnesota

In October, 1917, the writer began a survey[1] of all the schools in one of the representative counties in Minnesota, for the purpose of locating all cases of subnormality. Such a survey necessitates, among other things, visiting every school in the county, selecting all children of questionable mentality, and testing them, individually, with the Binet-Simon tests.

The selection of retarded and subnormal children above the third grade is relatively easy, for teachers' estimates, previous school records, and repetition of grades without the excuse of a long continued absence, taken together form a fairly accurate basis. However, such methods are practically impossible in and below the third grade. This is due to several factors: first, to the inaccuracy of the school records for primary grades; second, to the frequent change of teachers, especially in the rural schools; third, to the fact that many children have been in school too short a time to become sufficiently retarded pedagogically to suggest a diagnosis of mental deficiency; and fourth, to the inability of many children, on entering school, to do themselves justice in the grade because of the strangeness of the new environment. Other difficulties, common to all grades, are also encountered in such a survey. Teachers frequently persuade themselves that since a child is sweet, docile and attractive, he must necessarily be bright; or, because he is doing good work in his grade, he is undoubtedly normal, regardless of the fact that he is two or three years older than his classmates. In one rural school, the teacher, on being asked about the work of twin girls in the third grade, said with great assurance, "Oh, those girls are very bright. They do good third grade work, and they are such nice children." The twins in question had a chronological age of twelve years, a fact entirely overlooked by the teacher in her estimate of their ability. Mentally, they were found to be 7 5-8 and 8 3-8 years, respectively, when given the Binet tests. Hence, in order to have an accurate survey, the need for an impersonal method of selecting children of doubtful mentality was imperative.

To supply this need, the writer has devised a series of group tests, to be used in the first, second and third grades, in determining which children should be given the individual Binet tests. Some of these group tests are adaptations of the original Binet-Simon tests; some have been taken from Dr. Kuhlmann's 1917 Revision; others from Dr. Terman's Revision; still others were suggested by tests standardized by the Bureau of Analysis and Investigation in New York, and the rest are original.

The Selection of Group Tests. The requirements of the group test are many, for not only must it possess the characteristics necessary to the individual test, but it must also satisfy various other demands. Simplicity is an important criterion in the selection of the group test; simplicity of material; of directions; of response; and of scoring.

Since a large number of children must be tested at one time, it is necessary that tests be selected in which the material used can be easily carried and quickly distributed. In order to facilitate handling, and to avoid confusion, the writer has arranged the material for the group tests in booklets. To each child, who has been in school less than one year, *i. e.*, who is in the first grade, and hence can read and write but little, a booklet, 7½ in. x 7½ in., is given, which contains material and blanks necessary for the fifteen tests designated as five, six and seven year tests. Children in the second grade receive similar booklets designed for six, seven and eight year tests; while third grade pupils get booklets for the seven, eight and nine year tests. Thus, all the responses of each child are kept together, and the confusion incurred by distributing and collecting loose sheets of paper every few minutes is avoided. All pages of the booklet are numbered, and definite spaces are alloted to each test. All possible elimination of disturbing elements is important, for in the primary grades the attention of the children is easily distracted.

Not only must the material for the tests be simple, but the directions to be followed must be clear and brief. Herein lies the greatest difficulty in selecting tests for a group. Frequently one feels confident that the directions for a certain test are perfectly clear and that they could not possibly be misunderstood, and yet, when the test is given to a group of children, he finds that the meaning is entirely lost. Good English must frequently be sacrificed, for the child in the primary grades is surprisingly limited in vocabulary.

Originally, in giving the directions for one of the six years tests, the writer said, "Make a cross in the LARGEST square." The result was puzzling, for the children crossed the smallest square as often as they did the largest. The test apparently was a failure for that age group. However, before discarding it, the writer decided to experiment a little to see if the difficulty could be discovered. Instead of having the children cross the square, they were asked to point to the largest square, whereupon one little girl tearfully informed the writer that she "didn't know what that meant." That solved the problem. From then on the children were instructed to "Make a cross in the BIGGEST square," and the success of the test was assured. Brevity is equally essential in giving directions to children, for their habits of thinking, of concentrating the attention, are not yet formed, and to listen, comprehendingly, to long directions is impossible for them. Novelty and change are so obviously important in maintaining the interest and stimulating the child to his best effort, that further discussion of the value of brevity is unnecessary.

In the selection of group tests, only those should be chosen which permit of but one correct response, and that response must have but one possible interpretation. It is true that some of the tests consist of several trials of the same sort; thus, three different series of five digits each, are read to the children, to be reproduced, but the child passes the test if he succeeds in reproducing any one of the series correctly. This is a necessary procedure, because it is almost impossible, in a group, to have perfect silence while the directions are being given; and so a single series might not be heard by all. Then, too, many children have poor auditory imagery, which fact makes it difficult for them to reproduce material presented orally. However, in all of the tests, a response is either entirely correct or it is entirely wrong. This "All or none" method of scoring is used in the writer's group tests, for the purpose of the series is not to discover how much better one child can perform a given task than another, but to see which children fail to perform it at all, *i. e.*, to sort out the subnormal children who need to be given the individual tests before a correct diagnosis can be made regarding their mentality. That the responses must permit of no variation in interpretation, due to the examiner's judgment, is undoubtedly as essential in group tests as in the individual tests.

Altogether, twenty-five tests have been selected which seem to fulfill the requirements mentioned above. These have been divided into groups of five, designated for convenience, as five, six, seven, eight and nine year tests. These have been given to 566 school children distributed according to age as shown in Table I.

Table I.

Age		*No. Cases.*
6	(i. e., 5 yrs. 6 mos. – – – 6 yrs. 5 mos. incl.)	74
7	(i. e., 6 yrs. 6 mos. – – – 7 yrs. 5 mos. incl.)	132
8	(i. e., 7 yrs. 6 mos. – – – 8 yrs. 5 mos. incl.)	155
9	(i. e., 8 yrs. 6 mos. – – – 9 yrs. 5 mos. incl.)	115
10	(i. e., 9 yrs. 6 mos. – – – 10 yrs. 5 mos. incl.)	90
	Total	566

GROUP TESTS

TESTS FOR YEAR V.

V, I. Formboard

Materials.

This is an adaptation of Goddard's [1] formboard. From heavy white cardboard the forms shown in black in Fig. I are cut. Black paper is pasted over the back of the openings to make them stand out more clearly, and then another sheet of white cardboard is pasted over all this to form the back. Ten little cards which exactly fit the ten forms, are placed in an envelope pasted on the back of the formcard.

Procedure.

The examiner holds up a booklet in front of the group of children to be tested, and says:

"Open these little books to this black and white card. In the envelope on the back you will find a number of little cards. Take these out of the envelope and put them in a pile on your desk. Now, there is one place in the black and white card into which each of these little cards will exactly fit, — like this. (Illustrate with square). See if you can find the place for each of the others."

Help may be given any child who has difficulty in finding the place for each card the *first* time. When all the children have the cards properly placed, have them remove them and again place them in a pile on the desk. Then proceed as follows:

"Now you know where each little card belongs. This time we are going to see who can put all the cards in their places *first*. Everyone wait until I say 'GO,' and then work as fast as you can until I say 'STOP.' Then put your hands in your laps AT ONCE, without touching a single card again. Is everybody ready? GO."

Allow 50 seconds before giving the signal to stop. Care must be taken to see that no child touches a card after he has been told to stop.

Scoring.

The test is passed if all ten cards are in their proper places at the end of 50 seconds. A plus or minus, indicating whether the child has passed or not, can be quickly marked on the blank page opposite the form card, by the examiner, before having the children replace the cards in the envelope.

[1] This survey was made under the direction of Dr. Fred Kuhlmann. Director of Research, Faribault, Minnesota.

[1] H. H. GODDARD. *The Training School.* Vol. IX, No. 4, June 1912, pp. 49-54.

V, 2. Counting 4 Circles. (Lowell)

Materials.

A card, 6 in. x 18 in., on which are printed four large circles about 2 inches apart, (See Fig. VII), is necessary, and also two wooden blocks or other small objects. Space for the child to make the crosses is provided at the top of page 2 of the booklet.

Procedure.

"When I say, 'Make a cross,' this is what I want you to make. (Make X on the blackboard). If I should say, 'Make a cross for each of these blocks, (holding up two blocks), how many crosses would you make? Yes, two,—one for this block, and one for that, (XX), see?

"On this large card are some circles. When I hold the card so you can see them, you are to count the circles TO YOURSELVES, and then make as many crosses up here at the top of your paper, (Indicate), as there are circles. Ready. Look." Expose the circles for 10 seconds.

Scoring.

The test is passed if four crosses are made.

V, 3. Copying Square. (Binet)[1]

Materials.

Use a large cardboard on which an eight inch square is printed in heavy lines.

Procedure.

Holding up the large square, and indicating the space in the booklet to be used, the examiner says:

"At this side of the big space, here, see how nicely you can make a square just like this one."

Hold the card so all can see it while they are drawing, and then when they have finished one, say:

"Now see if you can make a still better one at this other side of the big space."

Scoring.

The test is passed if one of the two squares drawn is as good as those in the score card used by Binet.

V, 4. Discriminating Colors. (Binet)

Since it is impossible to have children in a group test name colors, the following adaptation of Binet's color test was used. It necessarily makes a different test from the original, since the directions must be comprehended and followed exactly.

Materials.

One-half inch squares of the colors red, yellow, blue and green, are pasted on page 2 of the booklet in the order named.

Procedure.

"Look at the colors in the middle of this page. Listen carefully, and then do just as I ask you:

1. Make a cross *in* the *yellow* square, like this.
2. Draw a line *through* the *green* square, like this.
3. Now make a cross *above* the *red* square, like this.
4. Draw a line *under* the *blue* square, like this.

1 Annee psychologique (1911) vol. 17, pp. 145–201.

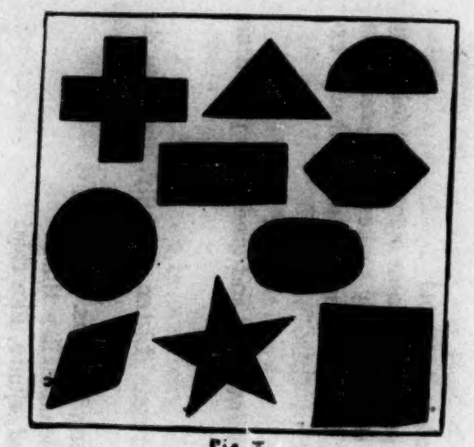


Fig. I.

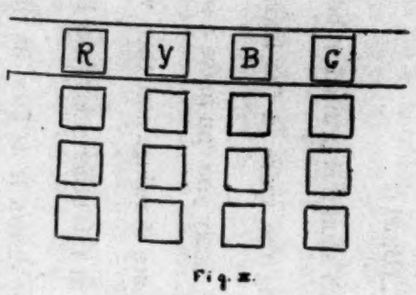


Fig. II.

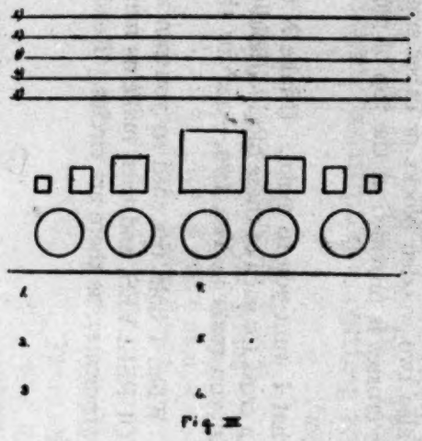


Fig. III

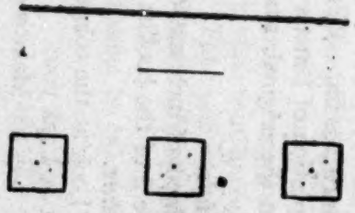


Fig. III.

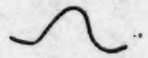


Fig. V.

Fig. VI.

Fig. VIII.

Fig. IX.

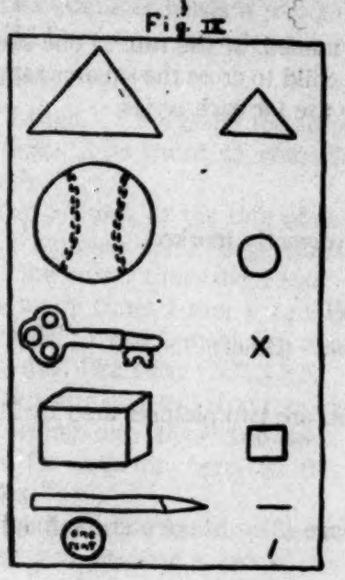
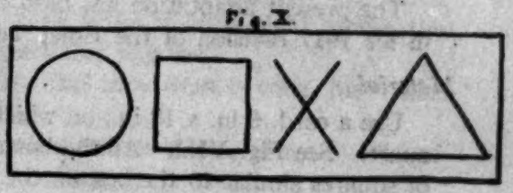
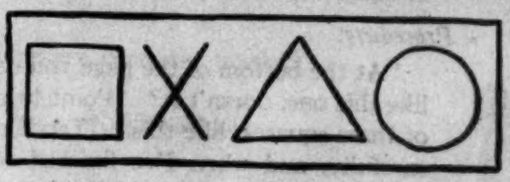
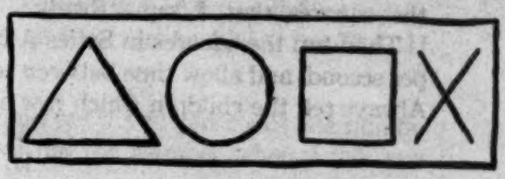
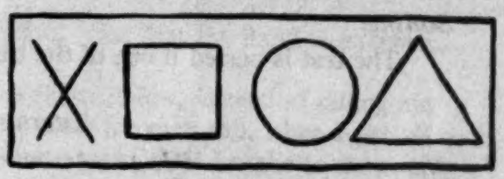


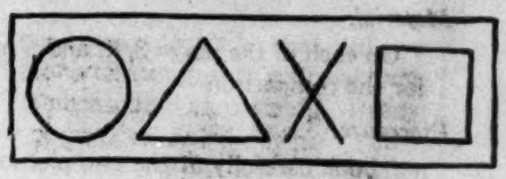
Fig. X.











In order that the child shall understand what is wanted, draw a square on the board each time, and do what you tell the child to do. Give directions slowly, and repeat each.

Scoring.

The test is passed if all four colors are correctly marked.

V, 5. Irregular Tapping. (Pintner;[1] Kuhlmann.[2])

The present adaptation has been made from the form used by Dr. Kuhlmann in his 1917 revision of the Binet Tests.

Material.

Use a card, 6 in. x 18 in., on which are printed four squares, about two inches apart. See Fig. VIII. At the bottom of page 2 in the booklet, are three rows of squares similar to the one on the examiner's card. See Fig. II.

Procedure.

"At the bottom of the page you see three rows of squares. The first row looks like this one, doesn't it? (Point to card in hand). Now I am going to tap some of these squares, like this. (Tap the first two squares with a ruler). Watch me carefully, and when I've finished tapping, put a cross, like this, X, in just the squares that I tap. Ready. Watch."

Then tap the squares in Series A in the order named, at the rate of one square per second, and allow time between series for the child to cross the squares tapped. Always tell the children which row of squares to use for each series.

A) 1–2–4
B) 1–3–4
C) 2–3–4

Scoring.

The test is passed if one of the three rows is correctly marked.

Tests for Year VI.

VI, 1. Aesthetic Comparison. (Binet)

Material.

On each of the pages 3, 4, and 5 of the booklet, are two pictures used by Binet for the comparison.

Procedure.

"Look carefully at the two pictures on this page (3). Make a cross under the prettier one of the two."

Give the same instructions for each of the pages 4 and 5.

Scoring.

The test is passed if the prettier face is indicated in two of the three series.

---

1 Rudolph Pintner and Donald G. Paterson. *A Scale of Performance Tests*, 1917.

2 Fred Kuhlmann. *The Measurement of Mental Development*, A School Publication, —Faribault—1917.

VI, 2. Mutilated Pictures. (Binet)

Materials.

The four pictures used by Binet are printed on four successive pages in the booklet, 6, 7, 8 and 9.

Procedure.

"Look at this picture. (page 6). Part of the face is gone. Let us see what part it is. You see the eyes. Look, the eyes are there. You see the nose. Yes, the nose is there, and look, the chin is there. Now what part is gone? (Children name mouth). Yes, the mouth is gone. Now make a cross right where the mouth should be, to show that that is the part that is gone.

"Look at the next picture. Do not tell me this time what is gone, just make a cross to show me what part is missing.

"Make a cross to show what is gone from the next picture. (Page 8) "And now show what is gone in the last picture." (Page 9).

Scoring.

The test is passed if the last three pictures are correctly marked.

VI, 3. Counting Irregular Taps. (Kuhlmann).

Materials.

The examiner needs a wooden block or the blunt end of a pencil for tapping. Ruled blanks are provided the children on page 10 of the booklet. See Fig. III.

Procedure.

"Listen. I am going to tap on the table and see if you can count the numbe of taps. You count to yourselves, and give me the number when I ask you. Ready."

Tap 5 times, at the rate of one tap per second. See that the hand is screened while tapping, by a large cardboard.

"How many times did I tap? Five, that is right. Now, instead of telling me how many times I tap, I want you to make a cross for each tap, when I get all through. I just tapped five times, so how many crosses shall I make to show it? Yes, five, like this: XXXXX.

"Sometimes I will stop tapping and then begin again. Don't let that fool you. You count only those you hear, and when I ask how many, you will make one cross for each tap, here, on this first line. (Indicate line each time). Ready. Listen."

Tap the following series in order, at the rate of one tap per second, for each cross, and pausing a second for each dash:

1)	X X - X X	(4)
2)	X - XX - X X X	(6)
3)	X X X - X	(4)
4)	X X - X X X - X	(6)
5)	X - X X - X - X	(5)

Scoring.

The test is passed if 3 out of the five series are correctly counted as indicated by the correct number of crosses.

VI, 4. Two Simultaneous Commands.

This test was suggested by Binet's test in which three simultaneous commands were used.

Materials.

One row of circles and a row of squares of different sizes are provided for the children on page 10 of the booklet. See Fig. III.

Procedure.

"See these squares and circles. Listen carefully, and see if you can remember what I am going to ask you to do with them, when I am all through telling you.

"Make a cross in the BIGGEST square.
Then draw a line in the FIRST circle."

"I will tell you once more, and then see if you can do both things. Listen:

"Make a cross in the BIGGEST square.
Draw a line in the FIRST circle."

Scoring.

The test is passed if both directions are correctly followed.

VI, 5. Perception of Sound. (Lowell)

Material.

A bean-bag, a baseball, a penny, a pencil, a key, and a wooden block are used. Have a table which can be concealed from the children and from which the above named objects can be dropped. A card picturing each of the objects with the respective symbols to be drawn by the children, is tacked up in view of the children. See Fig. X. Space for six drawings is provided at the bottom of page 10 in the booklet. See Fig. III.

Procedure.

"What is this?"

Hold up each of the objects and have the children name it.

"Now look at this card, (Fig. X), and tell me what each of these pictures is." (Have them name them all).

"Now, I am going to drop each one of these on the floor where you cannot see it, and I want you to listen to the noise each makes, and see if you can tell which of these I drop.

"If I should drop the bean-bag first, you would draw, in this first space, here, this. (Point to symbol at the side of the picture of the bean-bag); if I dropped the key next, you would make this X in the second space, here; if I dropped the penny next you would make this line in the third space, etc.

"Now shut your eyes, and see if you can tell from the sound which one of these I drop, and then look at the card to see what to draw. Then draw that in the first space. Ready, listen."

Drop objects in the following order:
block
bean-bag
penny
key
ball
pencil

Scoring.

The test is passed if five out of the six objects are named in the correct order.

TESTS FOR YEAR VII.

VII, 1. Number of Toes.

Because it is impossible to prevent a large group of children from counting their fingers if one were to use the original Binet test, the number of toes on each foot and on both feet together, has been substituted.

Procedure.

"I am going to ask you some questions, and I want you to write the answers, here at A) – – –, at B) – – – and at C) – – –. Ready:

A) How many toes on your left foot?
B) How many toes on your right foot?
C) How many toes altogether on both feet?"

Scoring.

Since most children of seven years can make figures up to 10, further use of crosses in answering questions is not permitted. The test is passed if the answers to all three questions are correct.

VII, 2. Memory Span. (Binet).

Procedure.

"I am going to read you some numbers. Listen carefully, and see if you can write them when I get through, just the way I read them."

Read the digits at the rate of one digit per second, without rhythm.

"Listen:

A) 6-5-2-8-1
B) 4-9-3-7-5
C) 2-8-6-1-9

Scoring.

The test is passed if the digits in one series are correct, even though their order be changed.

VII, 3. Copy Diamond. (Binet).

Material.

Use a card on which a large diamond, 9 in. x 18 in., has been drawn in heavy lines.

Procedure.

"At the left side of this big space, see how nicely you can draw a diamond just like this one." Hold the card so that all can see it as they draw. After they have finished drawing the first diamond, say:

"Now draw another one at the right side of the big space just like this one."

Scoring.

The test is passed if one of the two drawings is as good as those on the score card used by Binet.

VII, 4. Tying Bowknot. (Terman)[1]

Material.

Two pieces of tape, each about seven inches in length, are fastened to a stiff card pasted at the bottom of page 11, in the booklet.

Procedure.

"You know what kind of a knot this is, don't you? (Show them bowknot already tied). Yes, it is a bowknot. I want you to tie the same kind of a knot with the pieces of tape you have, just as quickly as possible."

Scoring.

The test is passed if a bowknot is correctly tied.

VII, 5. Geometrical Figures. (Lowell)

Material.

Use five cards, on which are printed, in a different order, on each card, a circle, a square, a cross, and a triangle. See Fig. IX.

Procedure.

"I shall show you some cards for ten seconds each. On each card you will find these forms. (Draw on blackboard). But the order in each case will be different. On one card a square will come first, and on another card the circle will come first, etc. You will notice carefully the order of the forms, and when I take the card away, you will draw what you saw in just the order it was on the card. Draw as quickly as possible. Ready. Look at the first card."

Allow not more than one minute for reproducing a card. Be sure all children are giving attention when you hold up a card.

Scoring.

The test is passed if three of the five series are correctly drawn.

TESTS FOR GROUP VIII.

VIII, 1. Ball and Field. (Terman).

Material.

The incomplete circle used by Terman to represent the baseball field, is printed in the booklet.

Procedure.

"Let us suppose that your baseball has been lost somewhere in this round field. You have no idea what part of the field it is in. All you know is that the ball is lost somewhere in this field. You are to mark out a path to show me how you would hunt for the ball so as to be sure not to miss it wherever it might be. If you draw just a line to the center from the gate, like this, — it would mean that you only looked that far and then stopped, — and the ball might be away over at the other side; you don't know where it is. Now begin at the gate, and with your pencil mark out a path to show how you would hunt for the ball so as to be sure not to miss it."

In giving an individual child this test, according to Dr. Terman's directions, the examiner watches what the child draws, and if he stops after drawing just a short path, the examiner asks him where he would go next if he hadn't yet found the ball. As such procedure is obviously impossible with a large group of children, the illustration is used in the directions.

1 L. M. TERMAN. *Measurement of Intelligence.* 1916—pp. 196–199.

Scoring.

The test is passed if the child shows a plan as good as those shown on Terman's score card for eight year old children.

VIII, 2. Writing from Dictation. (Binet)

Procedure.

"I want you to write something for me as nicely as you can on this long line. Write these words: 'See the little boy.' Be sure to write it all: 'See the little boy.'"

Scoring.

The test is passed if the sentence is written legibly enough to be easily read, and if no words are omitted. Unless a word is so incorrectly spelled as to make it unrecognizable, the error is not counted.

VIII, 3. Counting the value of Stamps. (Binet)

Material.

Use a cardboard on which have been pasted three one-cent stamps, in a row at the top, and three two-cent stamps in a row under these.

Procedure.

"Write the answer to this question on the short line, here." Hold the card before the children and say:

"How much will it cost to buy all these stamps? A green stamp is worth one cent, and a red stamp is worth two cents. Now, how much are they worth altogether?"

Scoring.

Test is passed if the answer, 9 cents, is given.

VIII, 4. Comprehension Test. (Kuhlmann).

Material.

Three one-inch squares are printed in the booklet.

Procedure.

Point to the parts as you mention them.

"See this first square. This is the *center* of the square. What is it? This is the *upper right corner* of the square. What is it? This is the *lower left corner*. What is it? This is the middle of the *left side of the square*. What is it?

"Now take your pencils, and draw a straight line from the center of the first (second and third) square to:

1) The upper left corner.
2) The middle of the upper side.
3) The lower right corner of the square."

Scoring.

The test is passed if two of the three lines are correctly drawn.

VIII, 5. Estimation of length of lines. (Kuhlmann).

Material.

On separate pages of the booklet are drawn three lines. The first line, drawn vertically in the center of the page, is three inches long; the second line, similarly drawn on another page, is two and one-fourth inches long; and the third line, drawn horizontally, and having two curves, is three inches long. See Fig. IV.

Procedure.

"I want to see how well you can judge the length of lines without a ruler or even your pencils to help you measure. Look at this line carefully, and then draw a line under it here, horizontally, *i. e.*, from left to right like this, (hold pencil where line is to be drawn), which is 1) Just as long and no longer than this line.

"Now look at this next line. Draw a line under it, here, horizontally, which is 2) Just twice as long as this one.

"Let's play that this is a piece of string. If I could take hold of the ends of it, and pull the string out straight, how long a piece would it be? Draw a straight line here, to show how long it would be," (Point to space under the curved line).

Scoring.

The test is passed if two of the three lines are drawn within one-half inch of the required length, i. e.,

1) 3 in.
2) 4½ in.
3) 3 in.

TESTS FOR GROUP IX.

IX, 1. Drinking Cup. (Bureau of Analysis, etc., N. Y.)[1]

Material.

A square of white paper, 7½ x 7½ in.

Procedure.

"I am going to show you how to make a drinking cup, and when I finish, I want you each to make one just like it. Watch, carefully.

1st: Fold the square through the center like this.

2nd: Take the right hand corner and fold it over to the opposite side so that these two edges are equal.

3rd: Next, take the other corner on the left side, and fold it over the same way so that the edges come straight across.

4th: Then, fold down this flap, and put it into this outer space to hold it.

5th: Now fold down the top flap for a cover, and you see how the cup looks. You take your square of paper now, and make one."

Scoring.

The test is passed if the five folds of the drinking cup are correctly made.

IX, 2. Maze. (Porteus)[2]

Material.

The maze used by Porteus for nine year children, is printed in the booklet.

Procedure.

"Here you have a maze. Some of these paths are open, so you can get through, and some are closed at the ends. See how quickly you can find the shortest way out, without getting into any of the closed paths. Start where you see the letter *S*, and mark out the very shortest path you can find in order to get out over here on the other side. Do not stop to erase anything. If you make a mistake, just draw a line across it and go on as fast as you can."

1 New York State Board of Charities, Bureau of Analysis and Investigation: *Eleven Mental Tests Standardized.* 1915. Eugenics and Social Welfare Bulletin No. V. p. 42.

2. S. D. PORTEUS. *Mental Tests for Feeble Minded.* Journal of Psycho-Aesthenics Vol. 19, No. 4, June 1915, pp. 200–213.

Scoring.

The test is passed if the shortest path is marked out in one minute, with not more than two errors. An error consists of a line, however, short, drawn out from the shortest path.

IX, 3. Tapping Squares. (Whipple).

Material.

A sheet of paper is used on which a rectangle has been printed. This rectangle is divided into 150 half-inch squares, arranged in ten rows of fifteen squares each.

Procedure.

"Please take your pencils and hold them firmly about half way up. Place your arm comfortably on the desk so that it may bend freely at the elbow. Now when I say, 'Ready,' you are to tap as many squares as you can, in order, without hitting any square twice, or missing a square, or touching any of the lines. Work until I say 'Stop.'

Scoring.

The test is passed if 50 squares are correctly tapped in the 30 seconds allowed.

IX, 4. Alphabet Test. (Kuhlmann)

Procedure

"Of course you all know the letters of the alphabet, a-b-c-d-e, etc. Who can tell me what is the first letter before 'C'? (B). Now who can tell me what is the sixth letter before 'X'? (R). Let us put this one on the board, R-S-T-U-V-W-*X*. Notice that W would be the first letter before X, V the second, U the third, T the fourth, S the fifth and R the sixth.

"I will give you 30 seconds to answer each one of these questions.

1)	What is the 3rd letter before	K	(H)
2)	What is the 5th letter before	W	(R)
3)	What is the 2nd letter before	F	(D)
4)	What is the 1st letter before	Q	(P)
5)	What is the 4th letter before	E	(A)

Always write the letter to be used as a basis on the board, so that the children will not misunderstand.

Scoring.

The test is passed if three of the five letters are correct.

IX, 5. Memory Test. (Kuhlmann).

Procedure.

"I am going to write five words on the board, and let you look at them for a few seconds. Then I will erase them and write five others and let you look at them, and so on with the other series. When I am through, I will give you the first word of a series, and you are to write down from memory, the other four words that go with it. Then I will give you the first word of the next list, and you will write the four that go with that one, etc. You must not write down any until I tell you to."

Let them look at each series for 30 seconds, then erase and take the next. Give them 30 seconds to recall each list. The lists are:

2 G. M. WHIPPLE. *Manual of Mental & Physical Tests.* 1910. pp. 100–115.

Silver	red	kitchen	blackboard	lawn
spoon	cow	stove	teacher	grass
table	fence	flour	desk	tree
bread	clover	water	slate	house
fork	dog	spoon	lesson	walk

Scoring.

The test is passed if the child gives three series without more than one error in each. An omission counts as an error.

Difficulties of Administering Group Tests. Many problems arise in giving group tests to the primary grades which, from their very nature, can never be entirely eliminated. Some of these difficulties are due to the distractibility of young children; to their inability to comprehend directions readily; to their undeveloped moral sense; to the attitude of the teacher; and to the apparent inhibition of response which the nervous child exhibits.

The inability of the young child to attend to set problems for any length of time is a serious difficulty. It necessitates devoting almost the entire first year of his school life to the formation of habits of work, of play, of thought and of living. Up to the time of his entrance into school, he is permitted to eat, sleep, and play largely at will, but with the beginning of his educational career he is expected to live according to a more or less definite plan, regardless of his own wishes or pleasure. When a stranger tries to test forty such children, before regular habits have been formed, the difficulties encountered may easily be imagined. Usually, the children are sufficiently shy and timid not to need disciplining, but sometimes that very timidity prevents them from being responsive at first. Therefore, their confidence must be won by the time the booklets are distributed and the examiner is ready to begin the testing. The fact that the children are not permitted to look through the pages except as they use them for the tests helps to keep up the interest. When the booklets are finally distributed, the directions for the first test given, the children attentively (?) awaiting the signal to begin work, and then a little voice in the corner confidentially informs one that "My sister has some gold beads just like yours," the examiner wonders if he can ever be sure of a child's attention.

The difficulty experienced by young children in comprehending directions has been discussed under the "Selection of Tests." Frequent illustrations, full and explicit directions with numerous repetitions, all help the child in understanding what he is to do.

However, ability to comprehend and follow directions accurately, is a test of intelligence, and too many or too great a variety of instructions may destroy the value of the test.

In order to stimulate the child to his best effort, one appeals to his competitive instinct, his sense of rivalry. This has its disadvantages, for the child in his eagerness to do better than his neighbor, or to get a word of praise from the examiner, resorts to methods "fair or foul." It is generally agreed that the young child is non-moral. The disapproval of the group has not been experienced sufficiently to make a lasting impression; and so the child slyly copies from his neighbor, or covertly writes on his desk the digits which are being read. To guard against such conditions as these, the writer has used various devices. Some have been mentioned under *Procedure*, such as having the child place his hands in his lap the instant he is given the signal to 'Stop'; covering his work with his hands as soon as he has finished the task assigned; or, in the digits test, place the pencils on the desks and leave them there until told to write. In one or two instances, where the attention had several times been called, without effect, to the fact that someone in the room was not playing fair, more drastic measures have been taken, such as speaking to the child by name, or changing his seat. However, as most children are very sensitive to criticism, such measures are seldom necessary.

Frequently a teacher, not knowing the object of the group testing, feels that unless the pupils respond well to the tests, her reputation as a teacher will be at stake. In her anxiety to have the children make a good showing, she offers suggestions which help them. For example, when trying out various tests for the nine year group, the writer used Terman's test of finding all the words that rhyme with "Day, mill and spring." In one third grade, the teacher felt that the children were not responding as readily as they should, so with much irritation she spoke to the class. "Why, children, you know how we have been finding words to go in the 'ING' family, so I don't see why you can't find others like 'Day' to go in the 'AY' family." Of course the children could find words after that, but the test, as such, had lost its value. To avoid such incidents, the examiner should explain to the teacher at the outset that the purpose of the work is to see what each child can do by himself, without any suggestion from her.

So far, no difficulty has been experienced in securing the proper attitude on the part of the child, for the interest and novelty of the tests counterbalance any temperamental disturbances. However, it has been evident in the cases of several nervous children that the excitement produced by the competition, novelty and mental effort has inhibited responses, especially where the test has a time limit. If a pupil happens to be an only child and has not yet learned to adapt himself to an environment in which there are many children, he may fail to make an average score in the group tests, though when tested individually he may test even above normal.

Reliability. Increase in percentage passing for successive chronological ages. The discriminative capacity of group as well as of individual tests, is measured by the increase in the percentage of children passing the tests from one age to the next. The greater the increase in the percentage passing any test from one age to the next, the higher is the discriminative capacity of that test. Table II compares first, the percentage passing at age for each test in the five age groups, with the percentage of children, chronologically a year older, passing each test; and second, it compares the per cent. increase from one chronological age to the next as found in the group tests, with the per cent. increase found by Kuhlmann [1] and Terman [2] where they have used the same tests in their revisions as individual tests. For example, 66 per cent. of the six year old children passed the aesthetic comparison test (VI,1) where given in a group, and 75 per cent. of the seven-year olds passed it, thus giving an increase from one age to the next of 9 per cent. Where Kuhlmann used it as an individual test he found an increase of 4 per cent. passing from 6 to 7 years, and Terman found an increase of 6 per cent. at the same ages.

Table II.

Group V	Yr. 6 (68 cases)	Yr. 7 (79 cases)	% Increase	Indiv. % Increase Kuhlmann	Indiv. % Increase Terman
1. Formboard	65	81	16	–	–
2. Counting 4 circles	63	81	18	1	–
3. Copying square	85	95	10	–	–
4. Discriminating Colors	78	81	3	–	9
5. Irreg. Tapping	70	85	15	–	–

1 These figures are taken from results in Dr. Kuhlmann's 1912 revision, in which the tests had not yet been correctly placed in the scale.

2 Terman's results as quoted are taken from the data published in 1917 on the Stanford Revision.

Group VI.	Yr. 6 (68 cases)	Yr. 7 (100 cases)			
1. Aesthetic Comparison	66	75	9	4	6
2. Mutilated Pictures	71	74	3	6	15
3. Irregular Taps	53	82	29	–	–
4. Two Commands	54	75	21	5	3
5. Perception of Sound	51	74	23	–	–
Group VII.	**Yr. 7 (59 cases)**	**Yr. 8 (136 cases)**			
1. No. of Toes	80	88	8	17	9
2. Memory Span	63	81	18	3	6
3. Copy Diamond	84	92	8	11	18
4. Tying Bowknot	68	73	5	–	16
5. Geom. Figures	61	76	15	–	–
Group VIII.	**Yr. 8 (137 cases)**	**Yr. 9 (115 cases)**			
1. Ball and Field	70	79	9	–	7
2. Dictation	78	87	9	–	8
3. Value of Stamps	70	79	9	20	17
4. Comprehension Test	72	78	6	–	–
5. Length of lines	69	81	12	–	–
Group IX.	**Yr. 9 (59 cases)**	**Yr. 10 (90 cases)**			
1. Drinking Cup	67	78	11	–	–
2. Maze	73	83	10	–	–
3. Tapping square	70	76	6	–	–
4. Alphabet Test	74	83	9	–	–
5. Memory Test	60	71	11	–	–

Unfortunately, there are no other similar group tests published with which the writer's percentages can be compared. However, there is about as much variation existing between the data from the two Revisions of the Binet Scale as between the individual and group results. In some instances, as in V, 4, Discriminating Colors, VI, 2, Mutilated Pictures, VII, 4, Tying Bow-knot, etc.—the tests show so little discriminative capacity as group tests, that the writer intends to find substitutes for them. Just why a test should show excellent discriminative capacity as an individual test and fail almost entirely when used in a group is at present a matter for conjecture to the writer. More data, however, will perhaps provide the explanation.

Considerable difference exists between the percentage of children passing at age in the group and in the individual tests. Dr. Kuhlmann has found that a much higher percentage of young children pass at age than of the older ones; thus, he found that 88 per cent. of the 4 year old children pass the 4 year tests, whereas only 54 per cent, of the twelve-year-olds pass the 12 year tests. However in the group tests, the writer finds as small a percentage of 6 and 7 year children passing 6 and 7 year tests, as of 9 year old children

passing the nine year tests. Table III compares the percentages passing at age as found by Dr. Kuhlmann [1] in the individual tests with those by the writer in the group tests. Thus, 78 per cent. of the six-year old children pass the six year individual tests, and only 60 per cent. of them pass the six year group tests.

Table III.

C. A.	*6* yrs.	*7* yrs.	*8* yrs.	*9* yrs.
Kuhlmann — Individual Tests	78	73	68	64
Lowell — Group Tests	60	69	72	68

The explanation of these differences in the six and seven year results may be found in the fact that a group test is largely a matter of comprehending and following directions presented orally, without any individual help for individual needs. If a child has not been used to doing things in a group, he will be easily distracted. Then, too, the younger the child the less capable he is of following instructions; but with each additional year of experience, and with an increasing mental age, the performance of tasks becomes easier. Therefore, the number of children passing a given test increases up to a certain point, with increasing chronological age. However, for ages eight and nine, the difference in percentage passing is possibly due to two causes, namely, the insufficient number of cases used in those age groups, and the questionable placing of some of the tests in the age groups. Further data are needed to corroborate this.

Increase in Number of Tests Passed for Successive Mental Ages.

If group tests are to be of value in selecting subnormal children, the number of tests passed in each age group by an individual must correlate with his age and intelligence. In the first three grades, used in the present survey, the chronological ages of the children ranged from six to eighteen years inclusive. The intelligence quotients of these children, determined after testing them with Dr. Kuhlmann's 1917 revision of the Binet tests, ranged from .40 — 1.30, though the largest number arrayed themselves between .70 — 1.00. In table IV, the average number of tests passed by the pupils of the various intelligence quotients, the average mental ages together with all the chronological ages found, are given for Grades I, II and III. Thus, in Grade I, chronologically six year old children, who had intelligence quotients from .70 — .79 inclusive, average

1 These figures were given me through the courtesy of Dr. Kuhlman, from his manuscript of *The Measurement of Mental Development, a further Extension and Revision of the Binet Simon Tests,* to be published soon.

4.7 years mentally and passed only 2.7 tests out of the age groups V, VI, VII; while those whose intelligence quotients were from .80 — .89. and whose mental age averaged 5.2 years, passed 4.3 tests, etc.

TABLE IV

Grade I (Groups V, VI, VII).

Chronological Age: [1]		6 yrs.	7 yrs.	8 yrs.	9 yrs.	11 yrs.
I. Q.						
.50– .59	Av. M. A.:	– –	4.0	– –	5.5	– –
	No. Tests passed:	– –	– –	– –	6.0	0 –
.60– .69	Av. M. A.:	– –	4.4	– –	– –	7.3
	No. Tests passed:	– –	1.9	– –	– –	7.9
.70– .79	Av. M. A.:	4.7	5.3	6.2	– –	– –
	No. Tests passed:	2.7	4.5	6.3	– –	– –
.80– .89	Av. M. A.:	5.3	5.9	6.6	7.5	– –
	No. Tests passed:	4.3	5.7	7.5	9.5	– –
.90– .99	Av. M. A.:	5.8	6.6	7.3	– –	– –
	No. Tests passed:	6.1	7.1	9.3	– –	– 0 –
1.00–1.09	Av. M. A.:	6.2	7.1	– –	– –	– –
	No. Tests passed:	8.3	9.8	– –	– –	– –
1.10–1.19	Av. M. A.:	6.5	– –	– –	– –	– –
	No. Tests passed:	9.0	– –	– –	– –	– –

Grade II (Groups VI, VII, VIII)

Chronological Age:		7 yrs.	8 yrs.	9 yrs.	10 yrs.	11 yrs.
I. Q.						
.60– .69	Av. M. A.:	– –	– –	5.5	– –	6.8
	No. Tests passed:	– –	– –	7.0	– –	11.0
.70– .79	Av. M. A.:	– –	6.3	6.4	7.3	8.2
	No. Tests passed:	– –	7.4	8.8	11.5	12.0
.80– .89	Av. M. A.:	6.5	6.9	7.6	7.9	– –
	No. Tests passed:	9.0	10.8	11.9	12.0	– –
.90– .99	Av. M. A.:	7.2	7.5	8.1	8.9	– –
	No. Tests passed:	10.0	11.0	12.3	13.0	– –
1.00–1.09	Av. M. A.:	7.8	8.7	8.9	– –	– –
	No. Tests passed:	11.0	12.6	13.0	– –	– –
1.10–1.19	Av. M. A.:	– –	9.3	– –	– –	– –
	No. Tests passed:	– –	14.0	– –	– –	– –
1.20–1.29	Av. M. A.:	– –	– –	– –	– –	– –
	No. Tests passed:	– –	– –	– –	– –	– –

Since, as has been stated before, a group test in the primary grades consists largely of understanding and obeying instructions, and since these are also the fundamental factors involved in the first three years of school work, it is evident that the more years a child spends in those grades, the easier it will be for him to follow directions, or in other words, the more group tests he will pass. For example, Table IV shows that in Grade I children who are chronologically 7 years and whose average mental age is 7.1 — 7.5 pass 9.8 tests out of Groups V, VI, and VII; whereas children in

[1]Chronological age 6 includes ages from 5 yrs. 6 mos, to 6 yrs. 5 mos. inclusive; chronological age 7 includes ages between 6 yrs. 6 mos. and 7 yrs. 5 mos., etc.

Grade III, (Groups VII, VIII, IX.)

Chronological Age:		8 yrs.	9 yrs.	10 yrs.	11 yrs.	12 yrs.	15 yrs.
I. Q.							
.40– .49	Av. M. A.:	– –	– –	– –	– –	– –	6.9
	No. Tests passed:	– –	– –	– –	– –	– –	6.3
.50– .59	Av. M. A.:	– –	– –	– –	– –	– –	8.1
	No. Tests passed:	– –	– –	– –	– –	– –	10.3
.60– .69	Av. M. A.:	– –	– –	– –	7.6	– –	9.4
	No. Tests passed:	– –	– –	– –	7.5	– –	13.3
.70– .79	Av. M. A.:	– –	7.0	7.7	8.3	8.8	–
	No. Tests passed:	– –	9.0	10.4	8.0	10.3	-
.80– .89	Av. M. A.:	6.8	7.8	8.5	8.8	9.6	–
	No. Tests passed:	8.5	10.6	12.2	9.0	11.3	–
.90– .99	Av. M. A.:	7.7	8.4	9.0	– –	– –	–
	No. Tests passed:	9 5	12.0	13.5	– –	– –	–
1.00–1.09	Av. M. A.:	8.7	– –	9.6	– –	– –	–
	No. Tests passed:	10.0	– –	14.0	– –	– –	–
1.10–1.19	Av. M. A.:	9.2	9.5	– –	– –	– –	–
	No. Tests passed:	12.0	15.3	– –	– –	– –	–

Grade II, with the same chronological age and average mental age, but who have been in school two years, pass 10 tests out of Groups VI, VII and VIII; and in Grade III children of the same chronological and average mental ages pass 12 tests out of Groups VII, VIII and IX. This increase in the number of tests passed, although the tests are of greater difficulty for each succeeding grade, seems to be accounted for by the additional years of school experience.

The use of the group tests in the survey has been to sort out the subnormal children, or those whose mentality is questionable, for further individual testing. If more complete standardization of the tests confirms the present results, it will be possible to make such a classification very readily. For example, upon giving the Group Test, V, VI and VII to a first grade, one could compare the number of tests passed by any child of a given chronological age with the number of tests passed by children of the same grade and chronological age, but with varying intelligence quotients, and when the numbers correspond the intelligence quotients can be approximated. Thus, if an eight year old child in the first grade passes only 6.3 tests, one may be fairly certain that his intelligence quotient will be somewhere between .70 — .79. Since this places him in the "borderline" group, he will undoubtedly need individual testing before his mental status can be correctly gauged.

COMMUNICATIONS AND DISCUSSIONS

A SCALE IN UNITED STATES HISTORY

The article in this JOURNAL for May, 1917, by Bell and McCollum on "A Study of the Attainments of Pupils in United States History" presents very obvious possibilities of making a much needed scale in that subject. To this end the writer has made certain calculations from the tabulated data in the article and has now re-issued the whole series of tests in convenient form with directions for giving the tests, scale values for each of the questions, directions for scoring, and standards for various school groups. The original data were obtained by Bell and McCollum with the use of the following test material in a large number of groups of students ranging from the fifth grade of the elementary school to the senior class in the University.

Test I. Give the reason for the historic importance of each of the ten representative dates: (1) 1861. (2) 1789. (3) 1620. (4) 1565. (5) 1898. (6) 1619. (7) 1783. (8) 1492. (9) 1776. (10) 1846.

Test II. Indicate for what each of these prominent men was celebrated: (1) John Burgoyne. (2) Alexander Hamilton. (3) Jefferson Davis. (4) Walter Raleigh (5) John C. Calhoun. (6) Cyrus H. McCormick. (7) George Dewey. (8) Sam Houston. (9) Roger Williams. (10) James Oglethorpe.

Test III. Mention the name of the man prominently connected with each of these historic events: (1) Captured Quebec during French and Indian War. (2) Discovered the North Pole. (3) Wrote the Declaration of Independence. (4) Invented the telephone. (5) Brought about the Missouri Compromise. (6) Captured the city of Mexico during the Mexican War. (7) Founded the Colony of Maryland. (8) Made a great speech against the English Stamp Act. (9) Was President of the United States during the Civil War. (10) Vetoed the re-chartering of the United States Bank.

Test IV. Define in a short sentence each of the following historic terms: (1) Second Continental Congress. (2) Lewis and Clark Expedition. (3) Articles of Confederation. (4) Sherman Anti-Trust Law. (5) Monroe Doctrine. (6) Fugitive Slave Law. (7) Dred Scott Decision. (8) Alien and Sedition Laws. (9) Nullification Ordinance of South Carolina. (10) Emancipation Proclamation.

Test V. Make a list of all the political parties that have arisen in the United States since the Revolution, and state one principle advocated by each.

Test VI. Indicate the great divisions or epochs of United States History.

Test VII. In the accompanying outline map of the United States draw the land boundaries of the United States at the close of the Revolution and indicate the different acquisitions of territory since that date. (Outline map 4 x 5 inches supplied).

The object of this report is to give some details of the derivation of the scale values and of the subsequent standard scales. The plan here adopted follows closely that employed by the writer in his

article "A Scale in Ancient History," published in the same number of the JOURNAL though there are some distinct variations and necessarily some limitations due to the fact that the original data was not obtainable and the published tables alone must be used. It will be observed from Table I of the Bell and McCollum article that forty-one groups of subjects were employed. The size of each group varied from eight to eighty-six and the total comprised 1476 individuals. Grades for these groups are presented in percentage of correct answers for each of the forty-three questions, except that there are no elementary school records for tests V, VI, and VII. Individual grades not being available, the nearest approximation to the distribution of those grades is the distribution of their group averages. Five calculations were made from the grades on representative questions and the Standard Deviations were found to be 17, 18, 18, 21, and 22. An average of these gives 19, a close approximation to the standard deviation for the various groups on any of the questions. In the following calculations, this is taken as the "Sigma" value. Its use is justified by rather bold assumptions but in spite of the apparent skew in the curve, its value is probably as well established as the nature of the data warrants.

The most representative set of grades available is that for 1425 Normal School, High School and Elementary School percentages shown in Table III, p. 270, of Bell and McCollum's study. Supplementing this by the weighted average grades of these people on tests V, VI, and VII taken from Table I, p. 264, achievement grades of 18%, 21%, and 38% respectively are obtained for this same group. These are shown in the "grade" column in Table I arranged as they were in the original test.

The median of these forty-three grades is found to be 28%. From this the "deviations" of the second column are obtained. Now, dividing each of these differentials by sigma (19), indices of the distance above or below the median of difficulty are found for each question. These indices are shown in the third column for each question, and bear reverse signs for the reason that the data is in terms of achievement while the subsequent scale values must be in terms of difficulty. To show this, the respective position on the distribution curve must be reversed. When the zero point on the base line is obtained, it must be a zero of difficulty of questions distributed along that base line, rather than a zero of ability to answer questions.

The method of approximating zero difficulty was by means of difficulty interpolation beyond the range in terms of range units of difficulty. This, of course, assumes that the group is fully representative. For instance, if the averages of the three extremes in each case are taken to determine range limits, they are found to be ten per cent. and seventy-nine per cent. respectively. This shows a range span of 69 per cent. The median deviations of these

TABLE I

Showing percental achievement grade, median deviations, index quotients of standard deviation distance on the base line, and derived scale values referred to an interpolated zero.

		Grade	Dev.	Index	Sc. Val.		Grade	Dev.	Index	Sc. Val.
I–	1	44	+16	−84	294	III–1	35	+7	−37	341
	2	17	−11	+58	436	2	75	+47	−247	131
	3	27	−1	+5	383	3	50	+22	−116	262
	4	17	−11	+58	436	4	15	−13	+68	446
	5	35	+7	−37	341	5	17	−11	+58	436
	6	17	−11	+58	436	6	15	−13	+68	446
	7	21	−7	+37	415	7	20	−8	+42	420
	8	89	+61	−321	57	8	26	−2	+11	389
	9	53	+25	−131	247	9	60	+32	−168	210
	10	22	−6	+32	410	10	12	−16	+84	462
II–	1	29	+1	−5	373	IV–1	21	−7	+37	415
	2	28	0	0	378	2	32	+4	−21	357
	3	45	+17	−89	289	3	25	−3	+16	394
	4	54	+26	−14	364	4	14	−14	+74	452
	5	39	+11	−58	320	5	35	+7	−37	341
	6	16	−12	+63	441	6	32	+4	−21	357
	7	43	+15	−79	299	7	21	−7	+37	415
	8	73	+45	−237	141	8	12	−16	+84	462
	9	33	+5	−26	352	9	7	−21	+81	459
	10	35	+7	−37	341	10	26	−2	+11	389
V–		18	−10	+53	431 (0)					
VI–		21	−7	+37	415 (0)					
VII–		38	+10	−53	325 (0)					

range limits are −18 and +51 respectively, which when divided by sigma, give indices of 94 and −268. Thus the span 69 per cent. is equal to the sum of these values, or 362 units. The upper limit of achievement, 79 per cent. is 21 per cent. from zero difficulty. Calculating from the above, zero is found to be 378 units from the median. Other spans of extreme and mean ranges were used in the calculation, but as the results differed only in fractions from this result, this was taken as a proper determination. Assuming, then, that zero difficulty is 378 units below the median of difficulty, and combining 378 algebraically with the values in column three, Table I, in this study, the final scale values of column four are obtained. In actual practice these are expressed in abridged form by moving the decimal point back to the left one place, except in the case of tests V, VI, and VII, where the grade is made up from a single measure, rather than from ten separate measures. This discrimination gives proper weight to these latter tests in the final score of the pupil.

In the scale as now made available in convenient form for use in schools these values are inserted beside each question for convenience of grading and the table of standards (Table II) is included.

TABLE II.

Showing total values for each test, achievement standards for four school groups, and sex differences for each of the seven tests.

	No.	Test I	TestII	TestIII	Test IV	Test VI	Test VII	Tot.
Value	1425	347	329	360	435	431	415	325	2669
University	75	153	179	126	174	142	265	146	1184
Normal School	207	149	176	171	152	134	203	107	1091
High School	668	90	121	106	97	109	141	181	745
El. School	550	37	63	46	25	X	X	X	171
Boys	579	128	141	133	109	134	145	117	907
Girls	846	111	115	101	91	100	146	85	740

These standards were obtained by turning the scale values back upon the percental achievements as distributed through Tables I, III and IV of Bell and McCollum's article.

The advantage of the scale over the original method of grading lies in the fact that it gives proper weight to the achievement of answering any one question. In the original study each question was graded as of equal difficulty with every other one. This was obviously not justified as a means of ranking pupils but was proper as a means of securing data on which to base a scale of difficulty of questions. In the standards here furnished, proper weighting is given to the answer for each question and the totals represent the probable achievement of each group. The method of grading is to credit the pupil with the score equal to the scale value of the question when it is accurately and adequately answered. In the case of partial answers the scale value should be appropriately apportioned according to the merit of the answer as measured by an absolute standard, rather than by one's idea of what a child of that particular grade might be expected to answer. Complete instructions are given with the test material itself.

In closing, the writer can announce the appearance in the near future of still another series of tests in Unites States History along quite divergent lines from this one. Instead of measuring the range of information in only a few aspects of the subject, the new series will more nearly cover the range of what history teaching should accomplish, and will likewise more nearly eliminate the subjective factor in grading.

The University of Texas. L. W. SACKETT.

CORRELATION BETWEEN TRANSLATION OF LATIN PROSE AND ETHICAL DISCRIMINATION[1]

The data for this study were obtained from twenty six juniors and seniors in the Ethical Culture High School, of New York City. They were given the following tests in the spring of 1918.

Latin Test.

Translate at sight the following text and write out your translation in good clean English, representing the thought of the author as carefully as possible. (The passage given was taken from Cicero and was so selected as to give a fair test of first-hand translation, but with familiar vocabulary).

Ethical Test.

Write out the answers to the following ethical problems according to your best judgment.

1. A throws a snowball through the school window. B sees him do it. When the teacher asks who broke the window, A remains silent. What should B do about it? Describe as many different points of view as possible and support the one you think right.
2. A is riding on a street car and the conductor fails to collect his fare. What should A do about it? Defend the position you take by giving as many reasons as you can think of.
3. A steals $5.00 and B steals $5,000. Which is the more guilty of doing wrong? Defend the position you take and give all the reasons you can.
4. A owes B $5.00. One day they meet at the cashier's window of a bank. As A turns away with his money he drops a $5.00 bill. Has B the right to keep the $5.00 which he sees A drop in payment of what A owes him? Defend the position you take with reasons.
5. Suppose that the day before their meeting at the bank B had asked A for the money on the street and A had refused to pay him and had abused him roughly for asking for the money. Under these circumstances would B have a right to keep the money he sees A drop. Defend the position you take with reasons.

---

[1]Pedagogical Studies in the Ethical Culture School No. 1. This study was made by the author while he was pedagogical physchologist in the Ethical Culture School, New York City.

No. of Pupil	Rank in Latin	Rank in Ethics	No. of Pupil	Rank in Latin	Rank in Ethics
1	1.0	22.5	14	13.0	13.0
2	2.5	17.5	15	13.0	22.5
3	2.5	21.5	16	17.5	1.0
4	5.0	5.0	17	17.5	2.5
5	5.0	8.5	18	17.5	5.0
6	5.0	13.0	19	17.5	24.5
7	8.5	17.5	20	21.0	2.5
8	8.5	21.5	21	21.0	8.5
9	8.5	24.5	22	21.0	17.5
10	8.5	26.0	23	23.5	8.5
11	13.0	8.5	24	23.5	17.5
12	13.0	13.0	25	15.0	5.0
13	13.0	13.0	26	26.0	13.0

Both sets of papers were gone over very carefully and sorted into ten piles each, from the best to the poorest, on the basis of keenness of discrimination manifested in the translation of the Latin and in the analysis of the ethical situations presented.

The following is a summary of the rank comparisons obtained:

The above comparisons yields a correlation of −.62, when calculated by the formula—Rho equals one minus six times the sum of the rank differences squared, divided by the number of cases times the quantity, the number of cases squared minus one. Since the form of distribution is approximately that of the normal probability curve, the correlation may be inferred from the obtained value of Rho by the use of table No. 36, p. 168, *Mental and Social Measurements*, by Thorndike.

If the formula R equals one minus six times the sum of the plus difference in ranks, divided by the number of cases squared minus one is used and the correlation value inferred from table No. 37 in Thorndike, the result is −.93.

ROBERT A. CUMMINS

Simpson College, Iowa.

TIME ASSIGNMENT AND RATE OF IMPROVEMENT IN HANDWRITING

There follow a description of the method and a statement of the results of an experiment conducted in grades six A, seven B, and seven A, in three of the white elementary schools of Guthrie, Okla., from January to May, 1915, covering a period of four and a half months, for the purpose of determining the effect upon improvement in handwriting of different amounts of time spent in practice.

Method.

The pupils of each school spent in class practice periods fifty (50), seventy-five (75), and one hundred (100) minutes per week, respectively. No attention was given to handwriting other than incidental attention in connection with all written work. There had always been a definite attempt on the part of all teachers concerned with instruction of classes under investigation to secure as good handwriting as possible in all written work, so that in this experiment no extra attention was given to that feature by the teachers of the fifty-minute and seventy-five minute groups.

The first week in January in the middle of the regular penmanship period three copies were made of a short verse which all the children of a room knew, and the best one was picked out; *i. e.* the one the teacher, without reference to any measuring scale, thought the best of the three. Then at the end of the training series, the middle of May, the same procedure was followed, using the same verse, the pupils not knowing at any time that any experiment was being carried out.

These samples of handwriting were all graded or measured the middle of May by the principals of these elementary schools, according to both the Freeman scale and the Ayres scale. One person graded the sixth A specimens of all three schools, both the January and the May specimens; one person, likewise, graded all the seventh B specimens, and one person, all the seventh A specimens. In this way we eliminated any discrepancy that might be due to different judges measuring or grading the work of pupils in some grades and different schools, and different judges measuring beginning and final tests of the same pupil. Beginning papers of the whole group were graded and then final papers, so that the judges, in grading final papers, would have no idea of any particular pupil's score on the beginning paper.

These results were tabulated, and from them the following tables were made, showing the amount of improvement, positive and negative, in each grade in each school, and the number of pupils making each amount of improvement. No attempt has been made to bring out all the possible interpretations of the data here presented, but only some of the general and tentative conclusions.

RESULTS

In considering results I will confine myself to measurement by the Freeman scale because it, by analyzing handwriting into its elements, makes possible more accurate measurement and diagnosis.

TABLE I

Amounts of improvement and number of cases of each amount in each grade and each school

Improvement	6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 10	Total cases
Seven A.		
School A.	0 0 0 0 0 0 2 0 3 4 0 1 0 0 1 0 0	11
School B.	0 0 0 0 1 2 5 4 5 5 0 1 3 1 2 0 1	30
School C.	0 0 0 0 1 2 1 0 2 6 2 4 2 1 1 0 0	22
Seven B.		
School A.	0 0 0 1 0 4 3 5 4 11 3 1 1 2 0 0 0	35
School B.	0 0 0 0 1 0 4 4 1 4 1 1 2 0 1 0 0	19
School C.	1 0 0 0 1 0 1 2 1 2 1 2 0 0 1 0 0	13
Six A.		
School A.	0 0 0 1 0 0 1 3 4 4 0 1 1 0 0 1 0	16
School B.	0 1 1 1 2 2 2 1 3 1 1 2 2 0 0 0 0	21
School C.	0 0 1 0 0 4 4 3 0 3 1 0 1 0 0 0 0	17
Totals(7A, 7B, 6A).		
School A.	0 0 0 2 0 4 6 8 11 19 3 3 2 2 1 1 0	62
School B.	0 1 1 1 4 4 11 11 7 12 3 4 6 1 3 0 1	70
School C.	1 0 1 1 1 6 6 5 3 12 4 6 4 1 1 0 0	52
Minutes per week—School A, 50; School B, 75; School C, 100.		

Comparing improvement of the different practice-period groups and showing per cent of whole group making different amounts of improvement we have

TABLE II.

Improvement	Zero & Neg.	1–3	4–6	7–10	Total
School A.	19.35%	61.29%	12.90%	6.45%	99.99%
School B.	31.43	42.85	18.57	7.14	99.99
School C.	30.77	38.46	26.92	3.84	99.99

The weighted average amounts of improvement of the three groups of pupils were as follows:—

School A, 2.37 (Median 2.52)
School B, 2.08 (Median 1.85)
School C, 2.13 (Median 2.70)

These figures show that the fifty-minute group made as a whole the most improvement. The hundred-minute group had, however, more cases of marked improvement. The negative improvement may be accounted for by the following:—1st. Some pupils, entering the schools that year for the first time had bad writing habits (especially bad movement) to be broken up and there was loss while this was being done. Shifting of school population was greatest in School B, and least in School C. 2nd. The stage of physical development of some pupils was undergoing rapid changes, new coordinations, etc., being established, so that there was a noticeable lack of mascular control. This is borne out by the fact that, while the six A group (ages 12-13) constituted but 29.34% of the three groups, 48.14% of the negative improvement was in this grade. 3rd. In some cases there seemed to be a lack of interest in penmanship improvement, and carelessness and inattention resulting from this, led to poorer work. School B had to contend with this condition the most, and School C, the least. If one eliminated negative improvement on the grounds that it (in the first two cases above) is a temporary condition, the weighted average improvement would then be as follows:— School A, 2.53
School B, 2.42
School C, 2.53

In view of the necessity for economy of time in education it is worth while that more extensive investigations, covering a longer period of time, and including many more pupils and a wider grade distribution of pupils, be made as to the effect of time assignment upon the rate of improvement. This experiment, covering four and a half months time, and including 184 pupils in grades six and seven indicates that (where ordinary care in handwriting is required in all written work) the amount of improvement is not proportional to the amount of time spent in practice; that the fifty-minute-per-week assignment to handwriting practice in these grades gives as good results as the seventy-five-minute or the hundred-minute assignment.

FOWLER D. BROOKS.
State Normal School, Mankato, Minn.

The Journal of Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except June to August by
WARWICK & YORK, INC.
HERSHEY, PA. BALTIMORE, MD.

Board of Editors

WILLIAM CHANDLER BAGLEY, CARL E SEASHORE,
J. CARLETON BELL, Managing Editor, GUY MONTROSE WHIPPLE.

——COLLABORATORS——

Sydney Alrutz,	H. H. Goddard,	Will S. Monroe,	L. William Stern
J. R. Angell,	Joseph Jastrow	E. C. Moore,	G. M. Stratton,
J. W. Baird,	G. E. Johnson,	M. V. O'Shea,	G. D. Strayer,
Bird T. Baldwin,	E. A. Kirkpatrick,	W. B. Pillsbury,	E. L. Thorndike,
W. Van Dyke Bingham,	E. H. Lindley,	S. H. Rowe,	E. B. Tichener,
E. F. Buchner,	Anna J. McKeag,	Sante de Sanctis,	J. B. Watson,
W. H. Burnham,	D. P. MacMillan,	E. G. Sanford,	W. H. Winch.
Ed. Claparede,	Adolf Meyer,	W. S. Small,	H. K. Wolfe,
S. S. Colvin,	J. B. Miner,	E. D. Starbuck,	A. H. Yoder.
W. F. Dearborn,			

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

NOTICE TO SUBSCRIBERS. Notify the publishers, Warwick & York, Inc., of any change of address; as copies lost in transit because of subscriber's removal to another address will not be duplicated

THE JOURNAL is issuing a series of Monograph Supplements for the publicaiion of more extensive researches in educational psychology. For details prospective contributors may apply to H. E. Buchholz, 10 East Center St., Baltimore, Md.

The price of the JOURNAL is $3.00 per year of nine monthly issues, payable in advance. Single issues are 50 cents each. Back volumes are $5.00 in original wrappers and $6.25 bound in library buckram. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc.

VOLUME X SEPTEMBER NUMBER 7

EDITORIAL

COMPOSITION SCALES FOR THOUGHT CONTENT

Undoubtedly the objective in most composition writing is good form of expression. The teacher's attention is fixed, not on what the pupil thinks, but on the way in which he expresses his thought. Some teachers, however, are shrewd enough to realize that the chief difficulty lies not in the mere form of expression, but in the fact that the pupil has no thoughts to express. These teachers insist that the real task of composition work is thought building. Instead of refining and polishing and deleting and rearranging, the composition teacher should take everything that is presented and call for more, – – always more. True, the pupil should be encouraged to recognize when he is thinking vigorously and richly, should be asked to discriminate between sturdy, luxuriant, copious thinking and lean, thin, halting paucity of ideas. The only way to do this, however, is to stimulate him to think, and that is chiefly an affair of subject matter, – – of content rather than of form. To teach composition,

then, you should let the pupil imagine that you care little or nothing about compositions as such, but that you are tremendously interested in some topic in which he is interested, and that you want to know everything about it that he can possibly find out and tell you.

To what extent do our composition scales help us to measure improvement in this sort of thing? Most of them are composed of detached and isolated paragraphs on topics in which a child would have only an artificial interest. But is it not possible to reduce to more exact terms any improvement in thought content? The teacher is well aware that in any class there are tremendous individual differences in thought content on any topic that may be proposed. The difficulty in trying to develop a scale to measure these differences is that one would have to have a different scale for each unit of subject matter that might be proposed for pupils' reflection. But perhaps the task is not so hopeless as it seems. If we should choose ten topics in which children are interested and on which some children think well, and should give each of them to a sufficiently large number of children, we might get such evenly distributed variations of thought content as would lend themselves to the divisions of a scale. On this basis we could construct ten scales based on representative types of thought content.

In dealing with any given pupil, John Smith, we could determine by a series of exercises his ability to think along each of these lines, and thus could construct a fairly accurate 'profile' of his thinking power. If he were high in some of these fields and low in others, the question would then arise whether it were better to try to develop thinking power along the lines in which he is now weakest, and thus attain what might be called 'all-round' thinking ability, or whether it would be more advantageous to cultivate intensively those fields in which he is already somewhat proficient. Further would come the question as to what type of exercise is best fitted to develop thinking power in a given field, what is John's individual learning capacity in the field in question, to what extent does his improvement in this field carry over into other fields, etc. There would undoubtedly be technical difficulties encountered in constructing such scales, but their value for diagnostic purposes and for studies in the art of teaching composition would be so great as to amply repay any effort that might be expended on them. With adequate control of the attitude of the writer toward the task (by asking him to place himself under definite conditions) important contributions might be made to the study of the development of the higher thought processes.

J. C. B.

NOTES AND NEWS

Dr. David Spence Hill, of the department of education, University of Illinois, has been elected president of the University of New Mexico.

Dr. W. W. Charters, dean of the school of education of the University of Illinois, has become professor of education and director of the research bureau for retail training at the Carnegie Institute of Technology. Dr. J. B. Miner, who has been acting director of this bureau, will continue with the bureau and will also give advanced courses in vocational psychology.

Dr. Grace E. Bird, psychologist at the Rhode Island Normal School, has been appointed professor of educational psychology at the Rhode Island State College. These two institutions are now affiliated for the purpose of giving degrees in education.

Dr. Clinton P. McCord, health director of the Albany, N. Y. schools, has been appointed head of the newly created bureau of health in the Pennsylvania State Department of Education.

The London Education Committee has recommended that Mr. Cyril Burt's position as psychologist should be made permanent.

Dr. George O. Ferguson, Jr., professor of psychology and education at Colgate University, has been appointed professor of education at the University of Virginia.

Professor Truman L. Kelley, who was in charge of the administration of the plan of the committee on education and special training for the rating and classification of soldiers in the S. A. T. C., has resumed his work in Teachers College, Columbia University.

At Teachers College, Columbia University, Franklin W. Johnson, principal of the University of Chicago high school has been appointed associate professor of education, Dr. Leta S. Hollingsworth has been promoted from instructor in educational psychology to assistant professor of education, and Dr. William A. McCall has been advanced from instructor in experimental education to assistant professor of education.

Dr. J. V. Breitwieser, professor of psychology and education at Colorado College, has been appointed associate professor of education in the University of California.

Dr. W. S. Foster, Assistant professor of psychology at Cornell University, has been appointed professor of psychology at the University of Minnesota.

Dr. Harry D. Kitson, instructor in psychology at the University of Chicago, has accepted the position in psychology at Indiana University made vacant by the election of Professor E. C. Lindley to the presidency of the University of Idaho.